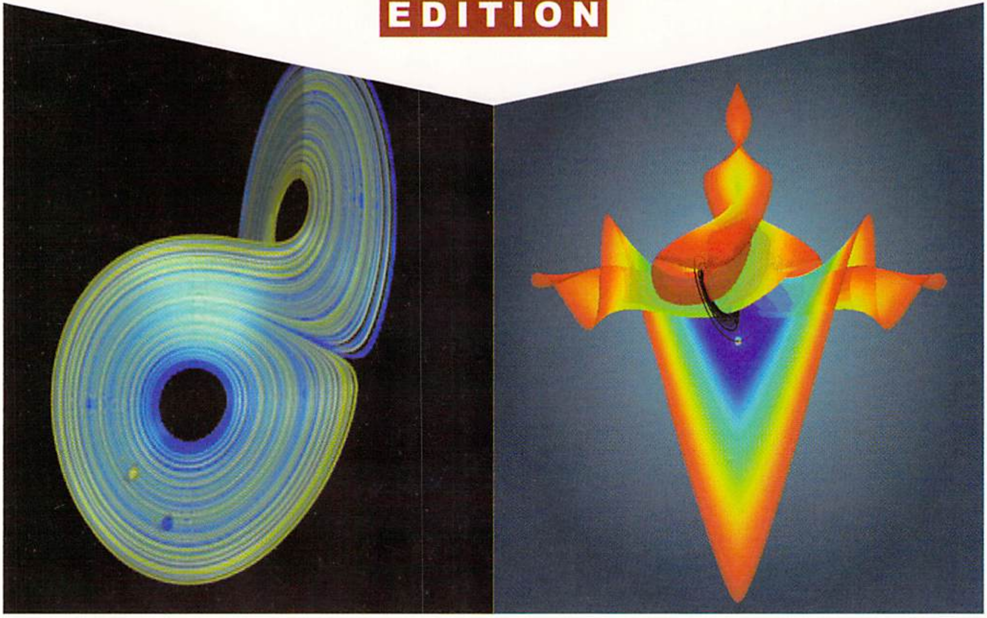
NEW AGE

MATHEMATICAL ANALYSIS

FIFTH Multi Colour EDITION



S C MALIK • SAVITA ARORA

NEW AGE INTERNATIONAL PUBLISHERS

MATHEMATICAL ANALYSIS

MATHEMATICAL ANALYSIS

FIFTH
Multi Colour
EDITION

S C MALIK
SAVITA ARORA
Former Senior Lecturer
Department of Mathematics
SGTB Khalsa College
University of Delhi
Delhi

NEW AGE
TECHNO
PRESS

An Imprint of

NEW AGE INTERNATIONAL (P) LIMITED, PUBLISHERS
LONDON • NEW DELHI • NAIROBI
Bangalore • Chennai • Cochin • Guwahati • Hyderabad • Kolkata • Lucknow • Mumbai
visit us at www.newagepublishers.com

Published by New Age International (P) Ltd., Publishers
First Edition: 1982
Fifth Multi Colour Edition: 2017

GLOBAL OFFICES

- **New Delhi** — **NEW AGE INTERNATIONAL (P) LIMITED, PUBLISHERS**
7/30 A, Daryaganj, New Delhi-110002, (INDIA)
Tel.: (011) 23253771, 23253472, **Telefax:** 23267437, 43551305
E-mail: contactus@newagepublishers.com • **Visit us at** www.newagepublishers.com
- **London** — **NEW AGE INTERNATIONAL (UK) LTD.**
27 Old Gloucester Street, London, WC1N 3AX, UK
E-mail: info@newacademicscience.co.uk • **Visit us at** www.newacademicscience.co.uk
- **Nairobi** — **NEW AGE GOLDEN (EAST AFRICA) LTD.**
Ground Floor, Westlands Arcade, Chiromo Road (Next to Naivas Supermarket)
Westlands, Nairobi, KENYA, **Tel.:** 00-254-713848772, 00-254-725700286
E-mail: kenya@newagepublishers.com

BRANCHES

- **Bangalore** — 37/10, 8th Cross (Near Hanuman Temple), Azad Nagar, Chamarajpet, Bangalore- 560 018
Tel.: (080) 26756823, **Telefax:** 26756820, **E-mail:** bangalore@newagepublishers.com
- **Chennai** — 26, Damodaran Street, T. Nagar, Chennai-600 017, **Tel.:** (044) 24353401, **Telefax:** 24351463
E-mail: chennai@newagepublishers.com
- **Cochin** — CC-39/1016, Carrier Station Road, Ernakulam South, Cochin-682 016
Tel.: (0484) 2377303, **Telefax:** 4051304, **E-mail:** cochin@newagepublishers.com
- **Guwahati** — Hemsen Complex, Mohd. Shah Road, Paltan Bazar, Near Starline Hotel, Guwahati-781 008
Tel.: (0361) 2513881, **Telefax:** 2543669
E-mail: guwahati@newagepublishers.com
- **Hyderabad** — 105, 1st Floor, Madhiray Kaveri Tower, 3-2-19, Azam Jahi Road, Near Kumar Theater, Nimboliadda Kachiguda, Hyderabad-500 027, **Tel.:** (040) 24652456, **Telefax:** 24652457
E-mail: hyderabad@newagepublishers.com
- **Kolkata** — RDB Chambers (Formerly Lotus Cinema) 106A, 1st Floor, S N Banerjee Road, Kolkata-700 014
Tel.: (033) 22273773, **Telefax:** 22275247
E-mail: kolkata@newagepublishers.com
- **Lucknow** — 16-A, Jopling Road, Lucknow-226 001, **Tel.:** (0522) 2209578, 4045297, **Telefax:** 2204098
E-mail: lucknow@newagepublishers.com
- **Mumbai** — 142C, Victor House, Ground Floor, N.M. Joshi Marg, Lower Parel, Mumbai-400 013
Tel.: (022) 24927869, **Telefax:** 24915415, **E-mail:** mumbai@newagepublishers.com
- **New Delhi** — 22, Golden House, Daryaganj, New Delhi-110 002, **Tel.:** (011) 23262368, 23262370
Telefax: 43551305, **E-mail:** sales@newagepublishers.com

ISBN: 978-93-859-2386-9
C-16-12-9897

Printed in India at IPP, Noida.
Typeset at In-house, Delhi.

NEW AGE INTERNATIONAL (P) LIMITED, PUBLISHERS
7/30 A, Daryaganj, New Delhi-110002
Visit us at **www.newagepublishers.com**
(CIN: U74899DL1966PTC004618)

Preface to the Fifth Edition

The book has been written in a fine thread of logic and provided a systematic approach to the development of the subject. I wish, I could do justice to the book in maintaining its superiority in presentation and clarity in writing its fifth edition. This effort is to pay my sincere regards to the departed soul, who not only taught me at the graduate and postgraduate level but also guided me like a father in every walk of life.

The revised edition of the book has been written to serve diverse set of students keeping in view their requirements as per the changed syllabi at the graduate and postgraduate level of some of the universities and to make the book more illustrative and self-contained. To this end a chapter on metric spaces containing various lucid examples, the topological framework—open and closed sets, convergence, completeness, compactness and connectedness, has been added. Differentiation under the integral sign and Beta-Gamma functions have been discussed in the chapter on double integrals and in Appendix I to meet the primary needs of the students at the graduate level.

To reinforce and solidify the understanding, some of the chapters and sections have been rearranged and several new exercises and solved examples have been incorporated. The section on limits inferior and superior of sequences is introduced and discussed in detail. Every care has been taken to explain and elucidate the different concepts so as to provide conceptual clarity to the readers.

I am grateful to Dr. S.C. Arora, Reader, Deptt. of Mathematics, University of Delhi, Delhi for his going through the chapters on Lebesgue Integral and Metric Spaces and making several useful suggestions. I am also thankful to Dr. Bansi Lal, Head, Deptt. of Mathematics, Kirori Mal College, Mrs. Shoba Rani, lecturer Vivekanand Mahila College, Delhi for pointing out various misprints and making fruitful suggestions. I am thankful to all my colleagues in the Department of Mathematics, especially to Mr. Satish Verma for the help he rendered continuously in preparing this revised edition. We hope this edition will prove to be useful for the students as well as teachers. Lastly we are thankful to M/s New Age International (P) Limited, Publishers for their keen interest in bringing out the revised edition of the book in an elegant form.

Suggestions and constructive criticisms for the improvement of the book are welcome.

AUTHOR

Preface to the First Edition

The book is intended to serve as a text for a course in real analysis that is usually taken up by the Honours and Postgraduate students of Indian universities, although its scope has been generally determined by the courses in real analysis prescribed by the University of Delhi. Professionals or those preparing for competitive examinations may also find in it much that is useful.

The purpose in writing this book has been to provide a development of the subject-matter which is well motivated, rigorous and at the same time not too pedantic. Most of the hard theorems which are either omitted or treated rather skimpily in many texts in advanced analysis have been proved with care. Some of them are ordinarily considered too difficult for a standard course in calculus but too elementary for a course of analysis. With the inclusion of those theorems, the book attempts to fill the gap and to make the transfer from elementary calculus to advanced course in analysis as smooth as possible.

Class work is often a joint effort between the teacher and the students, an effort which must adjust to the level of rigour demanded by the teacher and the capacity to assimilate, possessed by the students. The author hopes that the present book provides a course which will substantially meet both the requirements and will fulfill the need of a suitable text book.

The tone of the book is essentially classical but an attempt has been made to treat the subject matter on modern lines—some terminology has been updated and modern methods applied to smooth out and shorten certain classical techniques. The book starts with a quick review of the essential properties of rational numbers and using Dedekind's cut, the properties of real numbers are established. This foundation supports the subsequent chapters. The material on some of the topics—functions of several variables, uniform convergence of sequences and series of functions, line and surface integrals, double and triple integrals—is given with more details and with more examples and motivation than is usually done. No effort has been spared to present in a natural sequence the basic ideas of theory of Fourier series. The theory of surface integrals is generally paid less attention than it deserves. Special effort has been made to include all necessary details.

A large number of examples, taken mostly from the question papers of different universities and from my own class notes, have been properly graded and supplied with answers. Majority of them are straightforward; hints for harder ones are occasionally given. A sufficient number of them have been solved for the benefit of the readers. Short examples to illustrate the general theory or to show where it breaks down, follow every important principle. This and the remarks and notes added here and there should help in fixing the ideas better. Effort has also been made to give the proofs of a number of theorems in a somewhat modified form.

During the study of the subject and preparation of lecture notes for more than three decades of my University teaching (wherein lies the genesis of the present book) I have made use of the standard works of a large number of authors including Berman, Bromwich, Budak, Carslaw, Courant, Ferrar, Fomin, Gelbaum, Gibson, Goursat, Hyslop, Knopp, Landau, Olmsted, Nikolsky and Phillips. I have benefitted largely from them but I cannot resolve, how much is due to each. I owe much to them and I gladly take this opportunity to acknowledge my indebtedness to them. If due recognition for the authorship of any material is lacking, I extend my apology. Any such omission is unintentional.

I owe a special debt of gratitude of Professor U.N. Singh, Ex-Vice-Chancellor, University of Allahabad, who has been helping me all along with valuable suggestions.

I am deeply obliged to my reversed teacher Mr. Shanti Narayan, Ex-dean of Colleges, University of Delhi, for it is from him that I have derived my essential vision of the subject. My friends Dr. P.K. Jain, Reader, University of Delhi and Dr. S.C. Arora of Hans Raj College, Delhi, gave me the benefit of very fruitful discussions for which I am grateful. My sincere thanks are due to my colleagues in the Department of Mathematics in my college who helped me in various ways; and, in particular, to Dr. (Miss) Savita Arora who went through the first draft of the manuscript and suggested a number of improvements in the presentation of the subject.

Finally I may add that the author and also the publisher will welcome all suggestions for the improvement of the book.

S.C. Malik

Contents

1 Real Numbers

1. INTRODUCTION

In school algebra and arithmetic, we usually deal with two fundamental operations *viz. addition* and *multiplication* and their inverse operations, *subtraction* and *division* respectively. These operations are related to a certain class of 'numbers' which will be described more precisely in the following sections. The basic difference between 'elementary mathematics' and 'higher mathematics', which begins at the college level, is the introduction of the all important notion of **limit** which is very intimately related to the intuitive idea of *nearness* or *closeness* and which cannot be described in terms of the operations of addition and multiplication. The notion of limit comes into play in situations where one quantity depends on another varying quantity and we have to know the behaviour of the first when the second is *arbitrarily close* to a fixed given value. In order to illustrate our point in relation to a practical situation, consider the question of determining the velocity of the planet earth at a particular instant during its motion round the sun assuming that the path of its motion round the sun and its position on this path at any instant are known. We cannot determine the velocity of earth without taking recourse to the notion of limit and indeed we need the notion of limit even in defining the concept of 'velocity' of a moving object which is not moving with a uniform speed. The purpose of this illustration is simply to indicate that there are numerous situations where the methods of elementary algebra prove quite inadequate for the purpose of solving or even formulating a problem, and we are forced to evolve new concepts and methods. The notion of *limit* is one such concept.

For the proper understanding of the notion of limit and its importance, it is absolutely desirable that the reader should be familiar with the true nature and important properties of 'real numbers'. Starting with natural numbers, we shall briefly and intuitively introduce in the following sections the concept of *rational numbers* and *irrational numbers* which together form the system of *real numbers*, describing in the process the important properties possessed by these numbers. The branch of mathematics called *real analysis* deals with problems which are closely connected with the notion of 'limit' and some other notions, such as the operations of 'differentiation' and 'integration' which are directly dependent on the concept of limit when all these operations are confined to the domain of real numbers. It is difficult to say anything more precise at this stage. The interested reader will certainly have a clearer and precise understanding of this important branch of mathematics as he systematically studies this work.

1.1 Real Numbers

The system of real numbers has evolved as a result of a process of successive extensions of the system of **natural numbers** (*i.e.*, the positive whole numbers). It may be remarked here that the extension became absolutely inevitable as the science of Mathematics developed in the process of solving problems

from other fields. Natural numbers came into existence when man first learnt counting which can also be viewed as adding successively the number 1 to unity. If we add two natural numbers, we get a natural number—but the inverse operation of subtraction is not always possible if we limit ourselves to the domain of natural numbers only. For instance, there is no natural number which added to 8 will give us 3. In other words, 8 cannot be subtracted from 3 within the system of natural numbers. In order that the operation of subtraction (*i.e.*, operation inverse of addition) be also performed without any restriction, it became necessary to enlarge the system of natural numbers by introducing the *negative integers* and the number *zero*. Thus to every natural number n corresponds a unique negative integer designated $-n$ and called the additive inverse of n, and there is a number zero, written 0, such that $n + (-n) = 0$, and $n + 0 = n$ for every natural number n. Also n is the inverse of $-n$. The negative integers, the number 0, and the natural numbers (*i.e.*, the positive integers) together constitute what is known as the **system of integers**. Similarly, to make division always possible, zero being an exception, the concept of fractions, positive and negative, was introduced. Division by zero, however, cannot be defined in a meaningful and consistent manner. The system so extended, including integers and fractions both positive and negative, and the number zero, is called the **system of rational numbers**. Thus, every rational number can be represented in the form p/q where p and q are integers and $q \neq 0$.

We know that the result of performing any one of the four operations of arithmetic (division by zero being, of course, excluded) in respect of any two rational numbers is again a rational number. So long as mathematics was concerned with these four operations only, the system of rational numbers was sufficient for all purposes but the process of extracting roots of numbers (*e.g.*, square-root of 2, cube-root of 7, etc.), as also the desirability of giving a meaning to non-terminating and non-recurring decimals, necessitated a further extension of the number system. There were lengths which could not be measured in terms of rational numbers, for instance—the length of the diagonal of a square whose sides are of unit length, cannot be measured in terms of rational numbers. In fact, this is equivalent to saying that there is no rational number whose square is equal to 2. In order to be able to answer such questions, the system of rational numbers had to be further enlarged by introducing the so called irrational numbers. It is beyond the scope of this book to discuss systematically the definition of irrational numbers in terms of rational numbers. Numbers like $\sqrt{2}$, $\sqrt[3]{7}$, π (*i.e.*, the ratio of the circumference to the diameter of a circle) with which the reader is already familiar are examples of irrational numbers. Rational numbers and irrational numbers together constitute what is known as the **system of real numbers**.

Though the real number system cannot be extended in a way in which a rational number system is extended but it can be used to develop another system, called the **system of complex numbers**. But since real analysis is not concerned with complex numbers, we have nothing to do with complex numbers in this book.

For the sake of brevity and clarity of exposition, and because the notion of set is fundamental to all branches of mathematics, we start with the algebra of sets.

1.2 Sets

A set is a well defined collection of objects. In other words, an aggregate or class of objects having a specified property in common enables us to tell whether any given object belongs to it or not. The individual objects of the set are called *members* or *elements* of the set. Capital letters A, B, C, etc. are generally used to denote the *sets* while small letters a, b, c, etc. for *elements*. If x is a member of a set A, then we write $x \in A$ and read it as 'x belongs to A' or 'x is an element of A' or 'x is a member of A' or simply 'x is in A'. If x is not a member of A, then we write $x \notin A$ and read it as 'x does not belong to A'.

Some Typical Sets

N: The set of natural numbers,
I : The set of integers,
$\mathbf{I}^+$: The set of positive integers,
$\mathbf{I}^-$: The set of negative integers,
Q: The set of rational numbers,
R: The set of real numbers.

There are two methods which are in common use to denote a set.

(*i*) A set may be described by listing all its elements.

(*a*) Set S has elements a, b, c, then we write

$$S = \{a, b, c\}$$

(*b*) Set V of vowels in the English alphabets

$$V = \{a, e, i, o, u\}$$

(*ii*) A set may be described by means of a property which is common to all its elements.

(*a*) The set S of all elements x which have the property $P(x)$

$$S = \{x\text{: } P(x)\}$$

(*b*) The set B of natural numbers

$$B = \{n\text{: } n \in \mathbf{N}\}$$

Null Set. A set having no element. Sometimes the defining property of a set is such that no object can satisfy it, so that the set remains empty. Such a set is called a **null set**, an **empty set** or a **void set**, and is generally denoted by the Danish letter ϕ or { }. Thus

$$\phi = \{n\text{: } n \text{ is a natural number less than } 1\}$$

$$\phi = \{x\text{: } x \neq x\}$$

1.3 Equality of Sets

Two sets are said to be equal when they consist of exactly the same elements. Thus, sets P and Q are equal ($P = Q$) if every element of P is an element of Q, and every element of Q is also an element of P. Thus,

$$\{a, b, c\} = \{b, a, c\}$$

$$\{4, 5, 6, 7, 8, 9\} = \{n\text{: } 3 < n < 10,\ n \in N\}$$

It is to be noted that while writing a set, an element occurs only once but the order in which the elements of a set are written is immaterial.

A set is **finite** or **infinite** according to the number of element in it is finite or infinite.

1.4 Notation

$$\forall, \exists, \Rightarrow, \Leftrightarrow, \wedge, \vee, \sim$$

These symbols borrowed from mathematical logic help in a neat and brief exposition of the subject and so we shall describe them briefly here.

(*i*) $\forall$ stands for 'for all' or 'for every'.

The statement $x < y,\ \forall x \in S$ means x is less than y for all members of S, *i.e.*, all members of S are less than y.

(*ii*) $\exists$ stands for 'there exists'.

(*iii*) $\Rightarrow$ stands for 'implies that'.

If P and Q are two statements, then $P \Rightarrow Q$ means that the statement P implies the statement Q, *i.e.*, if P is true then Q is also true. Thus

$$x = 5 \Rightarrow x^2 = 25$$

$$AB \parallel CD \text{ and } CD \parallel EF \Rightarrow AB \parallel EF$$

If the statements P and Q are such that P implies Q and Q implies P, then we write

$P \Leftrightarrow Q$ (both ways implication)

Thus for real numbers x, y

$$xy = 0 \Leftrightarrow x = 0 \text{ or } y = 0$$

(*iv*) $\wedge$ stands for 'and'

$\vee$ stands for 'or'

The statement $P \wedge Q$ holds when both the statements P and Q hold, but the statement $R \vee S$ can hold when either R holds or S holds, *i.e.*, $R \vee S$ holds when at least one of R and S holds. Thus,

$$(x-3)(x-5) < 0 \Rightarrow x > 3 \wedge x < 5$$

$$x^2 = 1 \Rightarrow x = 1 \vee x = -1$$

(*v*) Negation $\sim$ stands for 'not'.

If P is a statement then $\sim P$ is negation of P.

In other words, $\sim P$ denotes 'not P'.

Thus when P holds, $\sim P$ cannot hold and *vice versa*.

$P \wedge \sim P$ is always false, but $P \vee \sim P$ is always true.

1.5 Subsets

If A and B are two sets such that each member of A is also a member of B, *i.e.*, $x \in A \Rightarrow x \in B$, then A is called a **subset** of B (or is contained in B) and we write $A \subseteq B$.

This is sometimes expressed by saying that B is a **superset** of A (or contains A) and we write $B \supseteq A$.

Thus, if A is a subset of B, then there is no element in A which is not in B, *i.e.*, $y \notin B \Rightarrow y \notin A$. Consequently, the null set ϕ is a subset of every set and $A \subseteq A$, for every set A.

Thus, if $A \subseteq B$ and $B \subseteq A$, we write $A = B$

$A \subseteq B$ allows for the possibility that A and B might be equal. If A is a subset of B and is not equal to B, we say that A is a **proper subset** of B (or is properly contained in B) and we write $A \subset B$. Thus A is a proper subset of B if every member of A is a member of B and there exists at least one member of B which is not a member of A.

Two sets A and B are said to be **comparable** if either $A \supseteq B$ or $A \subseteq B$, otherwise they are not comparable.

1.6 Union and Intersection of Sets

Union. If A and B are two sets, then the set consisting of all those elements which belong to A or to B or to both, is called the union of A and B and is denoted by $A \cup B$.

Clearly

$$A \cup \phi = A,\ A \cup A = A \text{ and } A \cup B = B \cup A.$$

Intersection. If A and B are two sets, then the set consisting of all those elements which belong to both A and B is called the intersection of A and B and is denoted by $A \cap B$.

Clearly

$$A \cap \phi = \phi,\ A \cap A = A \text{ and } A \cap B = B \cap A.$$

Thus, $A \cap B$ consists of elements which are common to A and B.

Two sets A and B are said to be **disjoint** if they have no common element, *i.e.*, $A \cap B = \phi$.

1.7 Union and Intersection of an Arbitrary Family

The operations of forming unions and intersections are primarily binary operations, that is, each is a process which applies to a pair of sets and yields a third. We emphasize this by the use of parentheses to indicate the order in which the operations are to be performed, as in $(A_1 \cup A_2) \cup A_3$, where the parentheses direct us first to unite A_1 and A_2, and then to unite the result with A_3. Associativity makes it possible to dispense with parentheses in an expression like this and to write $A_1 \cup A_2 \cup A_3$, where we understand that these sets are to be united in any order and that the order in which the operations are performed is irrelevant. Similar remarks apply to $A_1 \cap A_2 \cap A_3$. Furthermore, if $\{A_1, A_2, ..., A_n\}$ is any **finite** class of sets, then we can form

$$A_1 \cup A_2 \cup ... \cup A_n \text{ and } A_1 \cap A_2 \cap ... \cap A_n$$

in much the same way without any ambiguity of meaning whatever. In order to shorten the notation, we let $I = \{1, 2, ..., n\}$ be the set of subscripts which index the set under consideration. I is called the **Index Set**. We then compress the symbols and write

$$\bigcup_{i \in I} A_i \text{ and } \bigcap_{i \in I} A_i \quad \text{or} \quad \bigcup_{i=1}^{n} A_i \text{ and } \bigcap_{i=1}^{n} A_i$$

It is often necessary to form unions and intersections of large (really large) class of sets. Let Λ be a set and $\{A_\lambda : \lambda \in \Lambda\}$ an entirely arbitrary class or family **F** of sets which contains a set A_λ for each λ in Λ. Then

$$\bigcup_{\lambda \in \Lambda} A_\lambda = \{x : x \in A_\lambda \text{ for at least one } \lambda \text{ in } \Lambda\}$$

and

$$\bigcap_{\lambda \in \Lambda} A_\lambda = \{x : x \in A_\lambda \text{ for every } \lambda \text{ in } \Lambda\}$$

define the union and intersection of an **arbitrary family 'F'**.

Λ is called the **Index set**.

In particular, if $\Lambda = \{1, 2, 3, ...\}$ be the set of all natural numbers, then the union and intersection are often written in the form

$$\bigcup_{i=1}^{\infty} A_i \text{ and } \bigcap_{i=1}^{\infty} A_i$$

or simply $\cup A_i$ and $\cap A_i$.

1.8 Universal Set

In any discussion of sets, all sets are usually assumed to be subsets of a set, called the **universal set** (usually denoted by U). In our present discussion, however, the set **R** of real numbers can serve as the universal set.

1.9 Difference Set, Complement of a Set

If A and B are two sets, then the set consisting of those elements of B which do not belong to A is called the **difference set** of A and B and is denoted by $B - A$.

If, however, A is a subset of B then $B - A$ is called the **complement of A in B** or complement of A with respect to B.

Complement of A in the universal set U is called the **complement** of A and is denoted by A^c.

1.10 Functions

Let A and B be two sets and let there be a rule which associates to each member x of A, a member y of B.

Such a rule or a correspondence f under which to each element x of the set A there corresponds exactly one element y of the set B is called a *mapping* or a *function*.

Symbolically we write $f: A \to B$, *i.e.*, f is a mapping or a function of A *into* B.

The set A is called the *Domain* of the function.

The set B contains all the elements which correspond to the elements of A and is called the *co-domain* of f.

The unique element of B which corresponds to an element x of A is called the image of x or the value of the function at x and is denoted by $f(x)$; x is called the preimage of $f(x)$. It may be observed that while every element of the domain finds its image in B there may be some elements in B which are not the image of any element of the domain A. The set of all those elements of the co-domain B which are the images of the elements of the domain A is called the *range set* of the function f. If the co-domain B of f itself is the range set of f then we say that f is a function from A *onto* B. If members of the domain set are denoted by x and those of range set y then $y = f(x)$ is the value of the function f at x. We call a function $f: A \to B$ to be *one-one* if two different elements in A always have two different images under f, *i.e.*, $x_1 \neq x_2 \Rightarrow f(x_1) \neq f(x_2)$, for all $x_1, x_2 \in A$

If $f: A \to B$ is both *one-one* and *onto*, then we can define its inverse mapping $f^{-1}: B \to A$ as follows:

For each y in B, we find a unique element x in A such that $f(x) = y$ (x exists and is unique, since f is one-one and onto). We then define x to be $f^{-1}(y)$. The equation $x = f^{-1}(y)$ is the result of solving $y = f(x)$ for x.

If $f: A \to B$ is both *one-one* and *onto* then we say that f is a *one to one correspondence* between A and B. In this case $f^{-1}: B \to A$ is also a *one to one correspondence* between B and A.

If $A_1 \subseteq A$, then its image $f(A_1)$ is a subset of B defined by

$$f(A_1) = \{f(x) \in B: x \in A_1\}.$$

Similarly, if B_1 is a subset of B, then its inverse image $f^{-1}(B_1)$ is a subset of A defined by

$$f^{-1}(B_1) = \{x \in A: f(x) \in B_1\}$$

A function f is called an *extension* of a function g (and g is called a *restriction* of f) if the domain of f contains the domain of g

and $f(x) = g(x)$ for each x in the domain of g.

Just as we can combine sets to get a new set, we can combine given functions to construct a new function in the following way:

If $f\colon A \to B$ and $g\colon B \to C$ are two given functions, we define the *composite function* $g\, o\, f\colon A \to C$ by

$$(g\ o\ f)(x) = g\big(f(x)\big), \text{ for every } x \in A.$$

The function $I\colon A \to A$ defined by $I(x) = x$ for every $x \in A$ is called the identity function on A.

If $g\, o f = f\, o\, g = I$, then

$$g = f^{-1} \text{ or } f = g^{-1}.$$

The main properties of the function $f\colon A \to B$ and its inverse images are as follows:

(*i*) $f(\phi) = \phi$, where ϕ is an empty set.

(*ii*) $f(A) \subseteq B$

(*iii*) If $A_1 \subseteq A_2$, then $f(A_1) \subseteq f(A_2)$

(*iv*) $f\left(\bigcup\limits_i A_i\right) = \bigcup\limits_i f(A_i)$

(*v*) $f\left(\bigcap\limits_i A_i\right) \subseteq \bigcap\limits_i f(A_i)$

(*vi*) $f^{-1}(\phi) = \phi,\ f^{-1}(B) = A$

(*vii*) $f^{-1}\left(\bigcup\limits_i B_i\right) = \bigcup\limits_i f^{-1}(B_i),\ f^{-1}\left(\bigcap\limits_i B_i\right) = \bigcap\limits_i f^{-1}(B_i)$

(*viii*) $f^{-1}(B^c) = \left(f^{-1}(B)\right)^c$

1.11 Equivalent Sets

Two sets A and B are said to be equivalent (written as $A \sim B$) if there exists a one to one correspondence between their elements.

Let

$$A = \{a, e, i, o, u\} \text{ and } B = \{1, 2, 3, 4, 5\}$$

Then A is equivalent to B and the one-to-one correspondence can be seen as

$$a \leftrightarrow 1,\ e \leftrightarrow 2,\ i \leftrightarrow 3,\ o \leftrightarrow 4,\ u \leftrightarrow 5$$

Each of the two sets A and B have five elements which is a definite finite number and we call such sets as *finite sets*. Thus, if the sets are finite and have equal number of elements it is easy to see the one-to-one correspondence.

The positive integers are adequate for the purpose of counting any non-empty finite set; since all sets outside mathematics appear to be of this kind. But in mathematics we consider many sets which do not have a definite number of members. Such sets are called *infinite sets*.

The set **N** of all natural numbers, the set **I** of all integers, the set **Q** of rationals, the set **R** of reals, etc. are infinite sets.

The set **N** of naturals which is the same as the set of all positive integers seems to be larger than the set of all positive even integers $E = \{2, 4, 6, ...\}$, for **N** contains E as its proper subset. Does this mean that the set **N** has more elements than E? The answer is no. In dealing with infinite sets we must remember that the criterion for equivalent sets is whether there exists a one-to-one correspondence between these sets or not (irrespective of the fact which one is a proper subset of which). This function $f: \mathbf{N} \to E$ defined by $f(n) = 2n, n \in \mathbf{N}$ serves to establish a one-to-one correspondence between these sets. Thus, **N** is equivalent to E. Note that $\mathbf{N} \supset E$ but $\mathbf{N} \neq E$.

ILLUSTRATIONS

1. The set **N** of all natural numbers and the set S of all even integers are equivalent, a one-to-one correspondence is

$$1 \leftrightarrow 0, 2 \leftrightarrow 2, 3 \leftrightarrow -2, 4 \leftrightarrow 4, 5 \leftrightarrow -4, ...$$

2. The set **N** is equivalent to **I**,

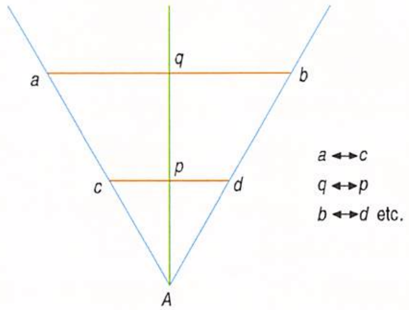
$$1 \leftrightarrow 0, 2 \leftrightarrow 1, 3 \leftrightarrow -1, 4 \leftrightarrow 2, 5 \leftrightarrow -2, ...$$

3. The set **N** and the set of all positive rationals are equivalent, the correspondence

$$1 \leftrightarrow 1, 2 \leftrightarrow \tfrac{1}{2}, 3 \to 2, 4 \to \tfrac{1}{3}, 5 \to 3,$$
$$6 \to \tfrac{1}{4}, 7 \to \tfrac{2}{3}, \text{ and so on}$$

has been set up, by adding up the numerator and denominator where sum is $2 : \frac{1}{1} = 1$, where sum is $3 : \frac{1}{2}, \frac{2}{1}$ where sum is $4 : \frac{1}{3}, \frac{2}{2}, \frac{3}{1}$, etc. and omitting those already listed.

4. The two closed intervals $[a, b]$ and $[c, d]$ are equivalent. The figure establishes a one-to-one correspondence between them.



1.12 Compositions

We shall be dealing mainly with number sets and so we define only two types of compositions in the sets.

An **Addition Composition** is defined in a set S if to each pair of members a, b of S there corresponds a member $a + b$ of S.

Similarly, a **Multiplication Composition** is defined in S if to each pair of members a, b of S there corresponds a member ab of S.

A set is said to possess an **algebraic structure** if the two compositions of Addition and Multiplication are defined in the set.

Subtraction and *Division* may be defined as inverse operations of addition and multiplication respectively.

Let $a, b \in S$.

Subtraction: $a - b$ may be expressed as $a + (-b)$ when $-b \in S$.

Division: The quotient a/b $(b \neq 0)$ may be put as $a \cdot 1/b$ or ab^{-1} when $1/b$ or $b^{-1} \in S$.

2. FIELD STRUCTURE AND ORDER STRUCTURE

2.1 Field Structure

A set S is said to be a **field** if two compositions of *Addition* and *Multiplication* be defined in it such that $\forall\ a, b, c \in S$ the following properties are satisfied.

A-1. Set S is closed for addition,

$$a, b \in S \Rightarrow a + b \in S$$

A-2. Addition is commutative,

$$a + b = b + a$$

A-3. Addition is associative,

$$(a + b) + c = a + (b + c)$$

A-4. Additive identity exists, *i.e.*, $\exists$ a member 0 in S such that

$$a + 0 = a$$

A-5. Additive inverse exists, *i.e.*, to each element $a \in S$ there exists an element $-a \in S$ such that

$$a + (-a) = 0$$

M-1. S is closed for multiplication,

$$a, b \in S \Rightarrow ab \in S$$

M-2. Multiplication is commutative,

$$ab = ba$$

M-3. Multiplication is associative,

$$(ab)c = a(bc)$$

M-4. Multiplicative identity exists, *i.e.*, $\exists$ a member 1 in S such that

$$a \cdot 1 = a$$

M-5. Multiplicative inverse exists, *i.e.*, to each $0 \neq a \in S$, $\exists$ an element $a^{-1} \in S$ such that

$$aa^{-1} = 1$$

A-M. Multiplication is distributive with respect to addition, *i.e.*,

$$a(b + c) = ab + ac$$

Thus, a set S has a **field** structure if it possesses the two compositions of addition and multiplication and satisfies the eleven properties listed above.

2.2 Order Structure

Ordinarily the order relation does not exist between the members of a general field, but as we are to deal with the field of real numbers, we can speak of one number being 'greater than' (or less than) the other.

A *field* S is an **ordered field** if it satisfies the following properties:

O-1. *Law of Trichotomy*: For any two elements $a, b \in S$, one and only one of the following is true.

$$a > b,\ a = b,\ b > a$$

O-2. *Transitivity*: $\forall\ a, b, c \in S$,

$$a > b \wedge b > c \Rightarrow a > c$$

O-3. *Compatibility of Order Relation with Addition Composition*:

$$\forall\ a, b, c \in S,$$
$$a > b \Rightarrow a + c > b + c$$

O-4. *Compatibility of Order Relation with Multiplication Composition*:

$$\forall\ a, b, c \in S,$$
$$a > b \wedge c > 0 \Rightarrow ac > bc$$

2.3 It may be seen that the set **Q** of rational numbers and the set **R** of real numbers are ordered fields while the set **N** of natural numbers and the set **I** of integers are not fields.

(*i*) The Set N of Natural Numbers

We begin the development of real numbers with the set **N** of natural numbers: 1, 2, 3, ... we could certainly list many properties of natural numbers, however, the following are taken as axioms:

P_1: $1 \in \mathbf{N}$; that is, **N** is a non-empty set and contains an element we designate as 1.

P_2: For each element $n \in \mathbf{N}$ there is a unique element $n_0 \in \mathbf{N}$ called the successor of n.

P_3: For each $n \in \mathbf{N}$, $n_0 \neq 1$; that is, 1 is not the successor of any element in **N**.

P_4: For each pair $n, m \in \mathbf{N}$ with $n \neq m$, $n_0 \neq m_0$, that is, distinct elements in **N** have distinct successors.

P_5: If $A \subseteq \mathbf{N}$, $1 \in A$ and $p \in A$ implies $p_0 \in A$, then $A = \mathbf{N}$.

These five axioms are called *Peano's postulate* and all known properties of natural numbers can be shown to be the consequences of these.

P_5 is called the principle of *Mathematical Induction*. From the principle of Mathematical Induction it follows that "Every non-empty subset of natural numbers has a first element", this is called the *well-ordering principle* for **N**.

The sum or product of any two members is easily seen to be a member of **N**, so that the set possesses two compositions of addition and multiplication, *i.e.*, the set **N** possesses an *algebraic structure*. However, it does not satisfy all the properties of a field (it does not possess additive identity, additive

inverse and multiplicative inverse) and hence the set of natural numbers is not a field. However, it has an *order structure* compatible with the *algebraic structure*.

(*ii*) The Set I of Integers

$$\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots$$

It may be easily seen that the set possesses an *algebraic structure* but does not satisfy all the properties of a field. (M-5–Multiplicative inverses do not exist.) Hence, the set of integers is not a field. However, it has an *order structure* compatible with the *algebraic structure*.

(*iii*) The Set Q of Rational Numbers

A rational number is of the form p/q, where p and q are integers and $q \neq 0$. Evidently, the set **Q** of rational numbers includes the set of integers.

A real number which is not rational (*i.e.*, cannot be expressed as p/q) is called an **irrational number**.

The set **R** of **real numbers** consists of rational and irrational numbers.

The sets **Q** and **R** satisfy all the properties (§ 2.1) of a field and are, therefore, called Fields. In addition to this, both these fields satisfy the four properties 0–1 to 0–4 (§ 2.2) of order, and hence form **ordered fields**.

2.4 Upto this stage we have discussed two properties—*the field property* and *the order structure property*. We have found that both the sets, the set **R** of real numbers and the set **Q** of rational numbers possess these properties. However, there is a property called the property of **completeness** which is possessed only by the set of real numbers and this distinguishes it from other sets of numbers. Let us, now, consider some notions and examples which will facilitate the study of that property.

2.5 **Example 1.** Show that there is no rational number whose square is 2.

■ Let, if possible, there exist a rational number p/q, where $q \neq 0$ and p, q are integers prime to each other (*i.e.* having no common factor) whose square is equal to 2,

i.e., $$(p/q)^2 = 2 \text{ or } p^2 = 2q^2 \qquad \ldots(1)$$

Now q is an integer and so is $2q^2$. Thus, p^2 is an integer divisible by 2. As such p must be divisible by 2, for otherwise p^2 would not be divisible by 2.

Let $p = 2m$, where m is an integer. Then, from (1),

$$2m^2 = q^2 \qquad \ldots(2)$$

Thus, it follows that q is also divisible by 2. Hence, p and q are both divisible by 2 which contradicts the hypothesis that p and q have no common factor. Thus, there exists no rational number whose square is 2.

Example 2. Show that $\sqrt{8}$ is not a rational number.

■ Let, if possible, $\sqrt{8}$ be the rational number p/q, where $q \neq 0$ and p, q are positive integers prime to each other, so that $\sqrt{8} = p/q$.

But $2 < \sqrt{8} < 3$

$\therefore$ $$2 < p/q < 3 \Rightarrow 2q < p < 3q \text{ or } 0 < p - 2q < q$$

Thus, $p - 2q$ is a positive integer less than q, so that

$\sqrt{8}\,(p - 2q)$ or $p/q\,(p - 2q)$ is not an integer.

But $\sqrt{8}(p-2q) = p/q(p-2q) = \dfrac{p^2}{q} - 2p$

$\Rightarrow \dfrac{p^2}{q^2}q - 2p = 8q - 2p$, which is an integer.

$\Rightarrow \sqrt{8}(p-2q)$ is an integer.

Thus, we arrive at a contradiction.

Hence, $\sqrt{8}$ is not a rational number.

Remark: We have considered $\sqrt{n}$ (n– not a perfect square), first when n was a prime and then n as a composite number. The procedures shown are typical and may be adopted under similar situations.

Ex. Show that there is no rational number whose square is
(*i*) 3, (*ii*) 5, (*iii*) 6.

2.6 Intervals – Open and Closed

A subset A of **R** is called an **interval** if A contains (*i*) at least two distinct elements and (*ii*) every element lies between any two members of A.

Open Interval. If a and b are two *real numbers* such that $a < b$, then the set

$$\{x : a < x < b\}$$

consisting of *all real numbers* between a and b (excluding a and b) is called an open interval and is denoted by $]a, b[$ or (a, b).

Closed Interval. The set

$$\{x : a \le x \le b\}$$

consisting of a, b and all real numbers lying between a and b is called a **closed interval** and is denoted by $[a, b]$.

Semi-closed or Semi-open Intervals.

$$]a, b] = \{x : a < x \le b\}$$

$$[a, b[= \{x : a \le x < b\}$$

The intervals are semi-closed or semi-open. The former is open at a and closed at b while the latter is closed at a and open at b.

3. BOUNDED AND UNBOUNDED SETS: SUPREMUM, INFIMUM

A subset S of real numbers is said to be **bounded above** if $\exists$ a real number K such that every member of S is less than or equal to K, *i.e.*,

$$x \le K, \quad \forall\, x \in S$$

The number K is called an **upper bound** of S. If no such number K exists, the set is said to be **unbounded above** or **not bounded above**.

The set S is said to be **bounded below** if $\exists$ a real number k such that every member of S is greater than or equal to k, *i.e.*,

$$k \le x, \ \forall \, x \in S$$

The number k is called a **lower bound** of S. If no such number k exists, the set is said to be **unbounded below** or **not bounded below.**

A set is said to be **bounded** if it is bounded above as well as below.

It may be seen that if a set has one upper bound, it has an infinite number of upper bounds. For if K is an upper bound of a set S then every number greater than K is also an upper bound of S. Thus every set S bounded above determines an infinite set—the set of its upper bounds. This set of upper bounds is bounded below in as much as every member of S is a lower bound thereof. Similarly, a set S bounded below determines an infinite set of its lower bounds, which is bounded above by the members of S.

A member G of a set S is called the **greatest** member of S if every member of S is less than or equal to G, *i.e.*,

(*i*) $G \in S$

(*ii*) $x \le G, \ \forall \, x \in S$

Similarly, a member g of the set is its **smallest** (or the least) member if every member of the set is greater than or equal to g.

Clearly, a set may or may not have the greatest or the least member but an upper (lower) bound of the set, if it is a member of the set, is its greatest (least) member. A finite set always has the greatest as well as the smallest member.

If the set of all upper bounds of a set S has the smallest member, say M, then M is called the **least upper bound** (l.u.b.) or the **supremum** of S.

Clearly, the supremum of a set S may or may not exist and in case it exists, it may or may not belong to S. The fact that supremum M is the smallest of all the upper bounds of S may be described by the following two properties:

(*i*) M is the upper bound of S, *i.e.*,

$$x \le M, \ \forall \, x \in S$$

(*ii*) No number less than M can be an upper bound of S, *i.e.*, for any positive number ε, however small, $\exists$ a number $y \in S$ such that

$$y > M - \varepsilon$$

Again it may be seen that *a set cannot have more than one supremum*. For, let if possible M and M' be two suprema of a set S, so that M and M' are both upper bounds of S.

Also M is the l.u.b. and M' is an upper bound of S.

$$\therefore \qquad M \le M' \qquad \text{...(1)}$$

Again M' is the l.u.b. and M is an upper bound of S.

$$\therefore \qquad M' \le M \qquad \text{...(2)}$$

From (1) and (2), it follows that $M = M'$.

If the set of all lower bounds of a set S has the greatest member, say m, then m is called the **greatest lower bound** (g.l.b.) or the **infimum** of S.

Like the supremum, the infimum of a set may or may not exist and it may or may not belong to S. It can be easily shown that a set cannot have more than one infimum.

The infimum m of a set S has the following two properties:

(*i*) m is the lower bound of S, *i.e.*,

$$m \leq x, \quad \forall\, x \in S$$

(*ii*) No number greater than m can be a lower bound of S, *i.e.*, for any positive number ε, however small, $\exists$ a number $z \in S$ such that

$$z < m + \varepsilon$$

ILLUSTRATIONS

1. The set **N** of natural numbers is bounded below but not bounded above. 1 is a lower bound.
2. The sets **I**, **Q** and **R** are not bounded.
3. Every finite set of numbers is bounded.
4. The set S_1 of all positive real numbers $S_1 = \{x : x > 0,\ x \in \mathbf{R}\}$ is not bounded above, but is bounded below. The infimum zero is not a member of the set S_1.
5. The infinite set $S_2 = \{x :\ 0 < x < 1,\ x \in \mathbf{R}\}$ is bounded with supremum 1 and infimum zero, both of which do not belong to S_2.
6. The infinite set $S_3 = \{x :\ 0 \leq x \leq 1,\ x \in \mathbf{Q}\}$ is bounded, with supremum 1 and infimum 0 both of which are members of S_3.
7. The set $S_4 = \left\{\dfrac{1}{n} : n \in \mathbf{N}\right\}$ is bounded. The supremum 1 belongs to S_4 while infimum 0 does not.
8. Each of the following intervals is bounded:

$$[a, b],\ ]a, b],\ [a, b[,\ ]a, b[.$$

Example 3. Prove that the greatest member of a set, if it exists, is the supremum (l.u.b.) of the set.

■ Let G be the greatest member of the set S.

Clearly,

$$x \leq G, \quad \forall\, x \in S$$

so that G is an upper bound of S.

Again no number less than G can be an upper bound of S, for if y be any number less than G, there exists at least one member g of S which is greater than y.

Thus, G is the least of all the upper bounds of S, *i.e.*, G is the supremum of S.

EXERCISE

1. Give examples of sets which are:
   (*i*) Bounded,
   (*ii*) Not bounded,
   (*iii*) Bounded below but not bounded above,
   (*iv*) Bounded above but not bounded below.

2. Find the infimum and the supremum of the following sets. Which of these belongs to the set?

(i) $[1, 3, 5, 7, 9]$ (ii) $\left\{-1, -\frac{1}{2}, -\frac{1}{3}, -\frac{1}{4}, \ldots\right\}$

(iii) $\left\{\frac{1}{n}; n \in \mathbf{N}\right\}$ (iv) $\left\{\frac{(-1)^n}{n}; n \in \mathbf{N}\right\}$

(v) $\left\{-2, -\frac{3}{2}, -\frac{4}{3}, -\frac{5}{4}, \ldots, -\frac{n+1}{n}, \ldots\right\}$ (vi) $\left\{1 + \frac{(-1)^n}{n}; n \in \mathbf{N}\right\}$

(vii) $]a, b[$ (viii) $[a, b[$.

3. Which of the sets in question 2 are bounded?
4. Find the smallest and the greatest members (if they exist) for sets in question 2.
5. Show that the greatest (or the smallest) member of a set, in case it exists, is unique.
6. Show that the smallest member of a set, if it exists, is the infimum of the set.
7. Is the converse of the solved example 3, true?
8. If $S \subseteq T \subseteq \mathbf{R}$, where $S \neq \phi$, then show that

   (i) If T is bounded above, then $\sup S \leq \sup T$;

   (ii) If T is bounded below, then $\inf T \leq \inf S$.

ANSWERS

2. (i) 1, 9; both (ii) $-1, 0$; infimum

   (iii) 0, 1; supremum (iv) $-1, \frac{1}{2}$; both

   (v) $-2, -1$; infimum (vi) $0, \frac{3}{2}$; both

   (vii) a, b; none (viii) a, b; infimum.

3. All sets are bounded.

4. (i) 1, 9 (ii) -1, does not exist

   (iii) does not exist, 1 (iv) $-1, \frac{1}{2}$

   (v) -2, does not exist (vi) $0, \frac{3}{2}$

   (vii) do not exist (viii) a, does not exist

4. COMPLETENESS IN THE SET OF REAL NUMBERS

We have seen that all the properties—the properties of an ordered field, described so far, are possessed by the two sets, the set of real numbers **R** and the set of rational numbers **Q**. We shall now state a property, the property of *completeness* (or *order-completeness*) which is possessed by **R** and not by **Q**. This property not only distinguishes **R** from **Q**, but together with the ordered field property, it characterises **R**, *i.e.*, the set of real numbers is the only set which is a *Complete Ordered Field*.

4.1 Order-Completeness in R

(**O-C**) *Every non-empty set of real numbers which is bounded above has the supremum* (*or the least upper bound*) *in* **R**.

In other words, the set of upper bounds of a non-empty set of real numbers bounded above has the smallest member.

If S is a set of real numbers which is bounded above, then by considering the set $T = \{x : -x \in S\}$ we may state the completeness property in the alternative form as:

Every non-empty set of real numbers which is bounded below has the infimum (or *g.l.b.*) *in* **R**. Or, equivalently the set of lower bounds of a non-empty set of real numbers bounded below has the greatest member.

We have thus completed the description of the set of real numbers as a **Complete Ordered Field**. We shall, however, show that the property of *completeness* does not hold good for the ordered field of rational numbers, *i.e.*, the ordered field **Q** of rationals is not order complete.

Theorem 1. *The set of rational numbers is not order-complete.*

To show that the set of rational numbers does not possess the property of completeness, it is suffice to show that there exists a non-empty set S of rationals (a subset of **Q**) which is bounded above but does not have a supremum in **Q**, *i.e.*, no rational number exists which can be the supremum of S.

Let S be the set (a subset of **Q**) of all those positive rational numbers whose square is less than 2.

i.e., $$S = \{x : x \in \mathbf{Q},\, x > 0 \wedge x^2 < 2\}$$

Since $1 \in S$, the set S is non-empty.

Clearly 2 is an upper bound of S, therefore, S is bounded above.

Thus, S is a non-empty set of rational numbers, bounded above. Let, if possible, the rational number K be its least upper bound. Clearly K is positive. Also by the law of trichotomy (0–1) which holds good in **Q**, one and only one of (*i*) $K^2 < 2$, (*ii*) $K^2 = 2$, (*iii*) $K^2 > 2$ holds.

(*i*) $K^2 < 2$. Let us consider the positive rational number

$$y = \frac{4 + 3K}{3 + 2K}$$

Then,

$$K - y = K - \frac{4 + 3K}{3 + 2K} = \frac{2\left(K^2 - 2\right)}{3 + 2K} < 0$$

$$\Rightarrow \qquad y > K \qquad \ldots(1)$$

Also,

$$2 - y^2 = 2 - \left(\frac{4 + 3K}{3 + 2K}\right)^2 = \frac{2 - K^2}{(3 + 2K)^2} > 0$$

$$\Rightarrow \qquad y^2 < 2 \Rightarrow y \in S \qquad \ldots(2)$$

Thus, the member y of S is greater than K, so that K cannot be an upper bound of S and hence, there is a contradiction.

(*ii*) $K^2 = 2$. We have already shown that there exists no rational number whose square is equal to 2. Thus, this case is not possible.

(*iii*) $K^2 > 2$. Considering the positive rational number y as defined in case (*i*), we may easily deduce from (1) and (2) respectively that

$$y < K \text{ and } y^2 > 2$$

Hence, there exists an upper bound y of S smaller than the least upper bound K, which is a contradiction.

Thus, none of the three possible cases holds. Hence, our supposition that a rational number K is the least upper bound of S is wrong. Thus, no rational number exists which can be the least upper bound of S.

Note: If we admit K in $\mathbf{R}$ and regard $\mathbf{S}$ as a set of real numbers then by the order completeness property, the supremum K of S exists in $\mathbf{R}$. Clearly $K > 0$ and

$$K^2 < 2 \Rightarrow y^2 < 2 \wedge y > K \Rightarrow K \neq \text{Sup } S$$

$$K^2 > 2 \Rightarrow y^2 > 2 \wedge y < K \Rightarrow K \neq \text{Sup } S$$

Thus by property 0–1, it follows that $K^2 = 2$, *i.e.*, the least upper bound K exists whose square is equal to 2. Further, since $K \notin \mathbf{Q}$, it follows that K is an irrational number. Similarly, it may be seen that there exist real numbers other than rational numbers whose squares are 2, 5, 7, ... etc. This establishes the *existence of irrational numbers*.

Ex. Show that the set of natural numbers is order-complete.

4.2 Archimedean Property of Real Numbers

The order-completeness property has important consequences, one of which is the Archimedean property of real numbers which we now proceed to prove.

Theorem 2. *The real number field is Archimedean, i.e., if a and b are any two positive real numbers then there exists a positive integer n such that na > b.*

Let a, b be any two positive real numbers and suppose, if possible, that for all positive integers $n(\in \mathbf{I}^+)$, $na \leq b$.

Thus, the set $S = \{na : n \in \mathbf{I}^+\}$ is bounded above, b being an upper bound. By the completeness property of the ordered-field of real numbers, set S must have the supremum M.

$$\therefore \quad na \leq M, \quad \forall\, n \in \mathbf{I}^+$$

$$\Rightarrow \quad (n+1)\,a \leq M, \quad \forall\, n \in \mathbf{I}^+$$

$$\Rightarrow \quad na \leq M - a, \quad \forall\, n \in \mathbf{I}^+$$

i.e., $M - a$ is an upper bound of S.

Thus, a number $M - a$ less than the supremum M (l.u.b.) is an upper bound of S, which is a contradiction and hence our supposition is wrong.

Hence, the theorem.

Corollary 1. If a be a positive real number and b, any real number then there exists a positive integer n such that $na > b$.

Corollary 2. For any positive real number a there exists a positive integer n such that $n > a$.

The result follows by considering the two positive real numbers 1 and a.

Corollary 3. For any $\varepsilon > 0$ there exists a positive integer n such that $1/n < \varepsilon$.

The result follows by taking $a = 1/\varepsilon$ in Corollary 2.

Corollary 4. If a be any real number then there exists a positive integer n such that $n > a$.

For $a \le 0$, any positive integer $n > a$, and for $a > 0$, result follows by Corollary 2.

Theorem 3. *Every open interval* $]\, a, b\, [$ *contains a rational number.*

Case I. If $0 < a < b$, by Corollary 3 there is a $m \in \mathbf{N}$ such that $1/m < (b-a)$. Let $A = \left\{ n \in \mathbf{N} \colon \dfrac{n}{m} > a \right\}$.

By Archimedean property $A \ne \phi$. Now by the well ordering principle for **N**, A has a first element say n_0 and so $n_0 - 1 \notin A$.

$$\therefore \qquad \frac{n_0 - 1}{m} \le a$$

$$\Rightarrow \qquad \frac{n_0}{m} \le a + \frac{1}{m} < a + (b - a)$$

$$\Rightarrow \qquad \frac{n_0}{m} < b. \qquad \text{But } n_0 \in A, \qquad \therefore \frac{n_0}{m} > a$$

Hence, there exists a rational number n_0/m in the open interval $]a, b[$.

Case II. If $a \le 0 < b$. Again by Corollary 3 there is a $n \in \mathbf{N}$ with $1/n < b$

Clearly $1/n \in \,]\, a, b\, [$.

Case III. $a < b \le 0$, then $0 \le -b < -a$. By the previous cases there is a rational number $q \in \,]-b, -a\, [$ and so the rational number $-q \in \,]\, a, b\, [$.

Corollary 5. Every open interval $]\, a, b\, [$ contains infinitely many rational numbers.

4.3 Dedekind's Form of Completeness Property

We now state the completeness property of real numbers in another form, due to Dedekind, which states:

If all the real numbers be divided into two non-empty classes L and U such that every member of L is less than every member of U, then there exists a unique real number, say α, such that every real number less than α belongs to L and every real number greater than α belongs to U.

Clearly, the two classes L and U so defined are disjoint and the number α itself belongs either to L or U. The property of real numbers referred to above is known as *Dedekind's property*. We may restate **Dedekind's Property**.

If L and U are two subsets of **R** such that

(*i*) $L \ne \phi, U \ne \phi$ (each class has at least one member),

(*ii*) $L \cup U = \mathbf{R}$ (every real number has a class)

(*iii*) Every member of L is less than every member of U, *i.e.*,

$$x \in L \wedge y \in U \Rightarrow x < y,$$

then either L has the greatest member or U has the smallest member.

4.4 Let us now prove the **equivalence** of the two forms of completeness.

(*a*) First we show that the *order completeness property of real numbers implies Dedekind's property*. The set **R** has the order completeness property, *i.e.*, every non-empty subset of **R** which is bounded above (below) has the Supremum (Infimum).

Let L, U be two subsets of $\mathbf{R}$ such that

(*i*) $L \neq \phi, U \neq \phi$,

(*ii*) $L \bigcup U = \mathbf{R}$, and

(*iii*) Every member of L is less than every member of U.

We have to show that either L has the greatest member or U has the smallest.

By (*iii*) the non-empty set L is bounded above. If L has the greatest member, it establishes the result. If L has no greatest member, then by the order completeness property, the set of its upper bounds, which coincides with U, has the smallest member. Thus either L has the greatest member or U has the smallest member.

(*b*) Let, now, $\mathbf{R}$ *satisfy the Dedekind's property*. We shall show that $\mathbf{R}$ also satisfies the order completeness property.

Let S be a non-empty set of real numbers bounded above, then we have to prove that S has the supremum.

Let L and U be two sets of real numbers defined by

$L = \{x: x$ is not an upper bound of $S\}$,

$U = \{x: x$ is an upper bound of $S\}$.

It may be easily seen that

(*i*) $L \neq \phi, U \neq \phi$,

(*ii*) $L \bigcup U = \mathbf{R}$, and

(*iii*) $x \in L \wedge y \in U \Rightarrow x < y$.

Then by Dedekind property, either L has the greatest member or U has the smallest member.

We shall show that L cannot have the greatest member.

Let, if possible, L has the greatest member, say ξ. Then

$$\xi \in L \Rightarrow \xi \text{ is not an upper bound of } S$$

$$\Rightarrow \quad \exists \text{ an } a \in S \text{ such that } \xi < a.$$

Now the real number $\dfrac{\xi + a}{2}$ is such that

$$\xi < \frac{\xi + a}{2} < a$$

Since $\dfrac{\xi + a}{2}$ is greater than the greatest member ξ of L,

$$\therefore \quad \frac{\xi + a}{2} \in U$$

$$\Rightarrow \quad \frac{\xi + a}{2} \text{ is an upper bound of } S. \qquad \text{...(1)}$$

Again, since $\dfrac{\xi + a}{2}$ is less than the member a of S,

$$\therefore \qquad \frac{\xi + a}{2} \in L$$

$$\Rightarrow \qquad \frac{\xi + a}{2} \text{ is not an upper bound of } S. \qquad \ldots(2)$$

Thus, we arrive at contradictory conclusions, and as such L has no greatest member. Thus, it follows that U, the set of upper bounds of S, has the smallest member, *i.e.*, the set S has the supremum.

We have thus proved the equivalence of Dedekind's and the order completeness property of **R**.

4.5 Explicit Statement of the Properties of the Set of Real Numbers as a Complete-Ordered Field

The set **R** of real numbers is a *complete-ordered field* because for arbitrary members, a, b, c of R, it satisfies the following conditions:

A-1. $a, b \in \mathbf{R} \Rightarrow a + b \in \mathbf{R}$

A-2. $a + b = b + a$

A-3. $(a + b) + c = a + (b + c)$

A-4. $\exists$ a member 0 in **R** such that
$a + 0 = a$

A-5. To each $a \in \mathbf{R}, \exists$ an element $-a \in R$ such that
$a + (-a) = 0$

M-1. $a, b \in \mathbf{R} \Rightarrow ab \in \mathbf{R}$

M-2. $ab = ba$

M-3. $(ab)c = a(bc)$

M-4. $\exists$ a member 1 in **R** such that
$a \cdot 1 = a$

M-5. To each $a \neq 0 \in \mathbf{R}, \exists$ an element $a^{-1} \in \mathbf{R}$ such that
$aa^{-1} = 1.$

A-M. $a(b + c) = ab + ac.$

O-1. For any two elements a, b of **R**, one and only one of the following is true:
$a > b,\ a = b,\ b > a$

O-2. $a > b \wedge b > c \Rightarrow a > c$

O-3. $a > b \Rightarrow a + c > b + c$

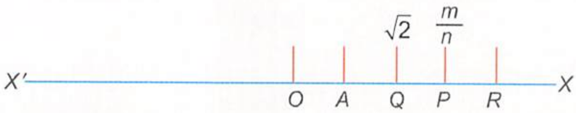
O-4. $a > b \wedge c > 0 \Rightarrow ac > bc$

OC. Every non-empty subset of **R** which is bounded above (below) has the supremum (infimum) in **R**.

4.6 Representation of Real Numbers as Points on a Straight Line

Points on a line can be used to represent real numbers. This geometrical representation of real numbers is sometimes very useful and suggestive especially to the beginner. But this should not stop us from giving the proper proof of a theorem which may otherwise seem to be obvious.

Let $X'X$ be a straight line. Mark two points O and A on it such that A is to the right of O.



The point O divides the line XX' into two parts; the part to the right of O containing A, may be called positive and that to the left of O as negative. Such a line for which positive and negative sides are fixed is called a *directed line*.

Let us consider the points O and A to represent rational numbers zero and 1 respectively, so that the distance OA is unity on a certain scale. To represent a rational number m/n $(n > 0)$, take a point P on the right of O if m is positive and to the left of O if m is negative, such that OP in m times the nth part of the unit length OA. Of course, the point P coincides with O if m is zero. The point P thus represents the rational number m/n. We may say that the rational number m/n corresponds to the point P or the point P corresponds to the rational number m/n. This way any rational number can be made to correspond to a point on the line. If points on the line corresponding to rational numbers be termed as *rational points*. We see that infinite number of rational points lie between any two different rational points, *i.e.*, *between any two rationals, there lie infinitely many rationals*.

Even though the rational points seem to cover a straight line very closely, there remain points on the line which are not rational. For example, the point Q on the line such that OQ is equal to the diagonal of the square with side OA does not correspond to any rational number. Also a point R such that OR which is a rational multiple of OQ, is also such a point. In fact there are infinitely many such points on the line. Hence, the set **Q** of rational numbers is insufficient to provide a complete picture of the straight line.

Such points on the line which are not rational, and which may be supposed to fill up the gaps between rational points are called *irrational points* and these correspond to irrational numbers. In fact, there is at least one irrational between two rationals. Thus like rationals, there are infinitely many irrationals. Hence, every real number can be represented on the directed line and there seem to be as many points on the directed line as the real numbers. The same fact is expressed by **Dedekind-Cantor Axiom** which states:

To every real number there corresponds a unique point on a directed line and to every point on a directed line there corresponds a unique real number.

In view of the order completeness property, the set of real numbers **R** does not have gaps of the kind **Q** has, and thus forms a continuous system. On account of this characteristic, the set **R** is called the *Arithmetical Continuum* and the set of points on a line as the *Geometrical Continuum*. In view of the above axiom, we see that there is a one-one correspondence between the two continuum and accordingly we may use the word *point* for *real number*, and the *real line* for the *directed line*. It is evident that between any two real numbers, there exist infinitely many real numbers both rational and irrational. This is the property of *denseness* of the real number system.

5. ABSOLUTE VALUE OF A REAL NUMBER

The *absolute value*, the *numerical value* or the *modulus* of a real number x, denoted by $|x|$, is defined as

$$|x| = \begin{cases} x, & \text{if } x \geq 0 \\ -x, & \text{if } x < 0 \end{cases}$$

Thus we always have

$$|x| \geq 0$$

Also by definition

$$|-x| = |x|$$

Some theorems which are immediate consequences of the definition will now follow:

Theorem 4. $|x| = \max(x, -x)$

Now $|x| = x \geq -x$, if $x \geq 0$

Also $|x| = -x > x$, if $x < 0$

Thus in either case $|x|$ is greater of the two numbers, x, $-x$, *i.e.*,

$$|x| = \max(x, -x).$$

Corollary 1. $|-x| = \max(-x, -(-x))$

$$= \max(-x, x) = |x|$$

$\therefore$ $|-x| = |x|$.

Corollary 2. $|x| = \max(x, -x) \geq x$

$\therefore$ $|x| \geq x$.

Theorem 5. $-|x| = \min(x, -x)$

Now

$$-|x| = -x < x, \text{ if } x > 0$$

Also

$$-|x| = -(-x) = x < -x, \text{ if } x < 0$$

Thus in either case $-|x|$ is smaller of the two numbers x and $-x$, *i.e.*,

$$-|x| = \min(x, -x)$$

Corollary. $-|x| = \min(x, -x) \leq x$.

$\therefore$ $-|x| \leq x$

Theorem 6. If $x, y \in \mathbf{R}$, then

(*i*) $|x|^2 = x^2 = |-x|^2$

(*ii*) $|xy| = |x| \cdot |y|$

(*iii*) $\left|\dfrac{x}{y}\right| = \dfrac{|x|}{|y|}$, provided $y \neq 0$

(*i*) For $x \geq 0$,

$$|x| = x \Rightarrow |x|^2 = x^2$$

For $x < 0$,

$$|x| = -x \Rightarrow |x|^2 = (-x)^2 = x^2$$

Thus in either case $|x|^2 = x^2$

Similarly, $|-x|^2 = (-x)^2 = x^2$

Hence, $|x|^2 = x^2 = |-x|^2$

(*ii*) $|xy|^2 = (xy)^2 = x^2y^2 = |x|^2 \cdot |y|^2 = (|x| \cdot |y|)^2$

$\therefore \qquad |xy| = \pm |x| \cdot |y|$

But since $|xy|$ and $|x| \cdot |y|$ are both non-negative, we take only the positive sign.

$\therefore \qquad |xy| = |x| \cdot |y|$

(*iii*) $\left|\dfrac{x}{y}\right|^2 = \left(\dfrac{x}{y}\right)^2 = \dfrac{x^2}{y^2} = \dfrac{|x|^2}{|y|^2} = \left(\dfrac{|x|}{|y|}\right)^2$

But since $\left|\dfrac{x}{y}\right|$ and $\dfrac{|x|}{|y|}$ are both non-negative, therefore taking positive square root of both sides, we have

$$\left|\frac{x}{y}\right| = \frac{|x|}{|y|}, \text{ when } y \neq 0.$$

Theorem 7. ***Triangle inequalities.*** *For all real numbers x, y show that*

(*i*) $|x + y| \leq |x| + |y|$, and

(*ii*) $|x - y| \geq ||x| - |y||$.

(*i*) *First Method*:

$$\begin{aligned} |x + y|^2 &= (x + y)^2 = x^2 + y^2 + 2xy \\ &\leq |x|^2 + |y|^2 + 2|x| \cdot |y| \qquad [\because\ xy \leq |xy| = |x| \cdot |y|] \\ &= (|x| + |y|)^2 \end{aligned}$$

Since $|x + y|$ and $|x| + |y|$ are both non-negative, therefore taking positive square roots on both sides, we have

$$|x + y| \leq |x| + |y|$$

Second Method: When $x + y \geq 0$.

$$|x+y| = x + y$$
$$\leq |x| + |y| \qquad [\because\ x \leq |x| \text{ and } y \leq |y|]$$

When $x + y < 0$,

$$|x+y| = -(x+y) = (-x) + (-y)$$
$$\leq |-x| + |-y| \qquad [\because\ -x \leq |-x| \text{ and } -y \leq |-y|]$$

But $|-x| = |x|$, $|-y| = |y|$

Thus in either case,

$$|x+y| \leq |x| + |y|.$$

(*ii*) *First Method*:

$$|x-y|^2 = (x-y)^2 = x^2 + y^2 - 2xy$$
$$\geq |x|^2 + |y^2| - 2|x| \cdot |y|$$
$$[\because\ -(xy) \geq -|xy| = -|x| \cdot |y|]$$
$$= (|x| - |y|)^2 = \big||x| - |y|\big|^2$$

Since $|x-y|$ and $\big||x| - |y|\big|$ are both non-negative, therefore taking the positive square root of both sides, we have

$$|x-y| \geq \big||x| - |y|\big|.$$

Second Method: Now

$$|x| = |(x-y) + y| \leq |x-y| + |y| \qquad \text{[by part } (i)]$$

$\therefore \qquad |x-y| \geq |x| - |y| \qquad ...(1)$

Again,

$$|y| = |(y-x) + x| \leq |y-x| + |x|$$

$\therefore \qquad |y-x| \geq |y| - |x| = -(|x| - |y|)$

But $\qquad |y-x| = |x-y|$

$\therefore \qquad |x-y| \geq -(|x| - |y|) \qquad ...(2)$

From (1) and (2),

$$|x-y| \geq \max\{|x| - |y|, -(|x| - |y|)\}$$
$$= \big||x| - |y|\big|$$

Hence, $\qquad |x-y| \geq \big||x| - |y|\big|$

Example 4. For real numbers $x, a, \varepsilon > 0$ show that

(*i*) $|x| < \varepsilon \Leftrightarrow -\varepsilon < x < \varepsilon$,

(*ii*) $|x - a| < \varepsilon \Leftrightarrow a - \varepsilon < x < a + \varepsilon$.

■ (*i*) $|x| = \max(x, -x) < \varepsilon$

$\Leftrightarrow \quad x < \varepsilon \wedge -x < \varepsilon$

$\Leftrightarrow \quad x < \varepsilon \wedge -\varepsilon < x$

$\Leftrightarrow \quad -\varepsilon < x < \varepsilon$

(*ii*) $|x - a| = \max\{(x - a), -(x - a)\} < \varepsilon$

$\Leftrightarrow \quad (x - a) < \varepsilon \wedge -(x - a) < \varepsilon$

$\Leftrightarrow \quad x < a + \varepsilon \wedge a - \varepsilon < x$

$\Leftrightarrow \quad a - \varepsilon < x < a + \varepsilon$

Example 5. Show that a set S of real numbers is bounded if there exists a real number $G > 0$ such that

$$|x| \le G, \quad \forall\, x \in S.$$

■ Suppose that S is bounded, therefore it is bounded both above and below. Let K be an upper bound and k, a lower bound for S.

On taking a real number $G = \max(|K|, |k| + 1)$, we have

$$K \le |K| \le G \text{ and}$$
$$-k \le |k| < |k| + 1 \le G \qquad i.e., \quad k > -G$$

This implies

$$-G < k \le x \le K \le G, \quad \forall\, x \in S$$

Hence, $\quad |x| \le G \quad \forall\, x \in S.$

The converse is trivial.

Ex. If a and b are real numbers, then show that

$$\max(a, b) = \frac{a + b + |a - b|}{2}$$

and

$$\min(a, b) = \frac{a + b - |a - b|}{2}$$

Example 6. If $a, b \in \mathbf{R}$ such that $a < b + \varepsilon$ for each $\varepsilon > 0$, then $a \le b$.

■ Suppose $a > b$. Then $a - b > 0$, so that

$$a < b + (a - b) \quad (\text{by taking } \varepsilon = a - b)$$

and so $a < a$

This is a contradiction. Hence our assumption $a > b$ must be false. Therefore $a \le b$.

Example 7. Let $a, b \in \mathbf{R}$. Show that if $a \le b + \frac{1}{n}$, for all $n \in \mathbf{N}$, then $a \le b$.

■ Assume $a \le b + \frac{1}{n}$, for all $n \in \mathbf{N}$ and $a > b$

Then $a - b > 0$ and by the Archimedean property, we have

$n_0(a-b) > 1$, for some $n_0 \in \mathbf{N}$

Then $a > b + \frac{1}{n_0}$, contrary to our assumption.

Example 8. If for any $\varepsilon > 0, |b-a| < \varepsilon$, then $b = a$

■ We have, for any $\varepsilon > 0, b < a + \varepsilon$ and $a - \varepsilon < b$. Since $b < a + \varepsilon$ for any $\varepsilon > 0$, it follows that $b \le a$. Since $a < b + \varepsilon$ for any $\varepsilon > 0$ this implies $a \le b$. Hence, $b = a$.

Example 9. If $a, b \in \mathbf{R}$ and $a < c$ *for each* $c > b$, then $a \le b$.

■ Assume that a and b satisfy the hypothesis but not the conclusion. Then $a > b$, and so there is a $c \in \mathbf{R}$ such that $a > c > b$. Now $c > b \Rightarrow a < c$ in contradiction to $a > c$.

EXERCISE

Prove the following (Qs.1-3):

1. $|x-y| \le |x| + |y|$.
2. $|x+y| \ge \big||x| - |y|\big|$.
3. (*i*) $\sqrt{x^2+y^2} \le |x| + |y|$,

   (*ii*) $\sqrt{|x+y|} \le \sqrt{|x|} + \sqrt{|y|}$.
4. If $x_1, x_2, x_3, \ldots, x_n$ are real numbers, then show that

   (*i*) $|x_1 + x_2 + \ldots + x_n| \le |x_1| + |x_2| + \ldots + |x_n|$

   (*ii*) $|x_1 x_2 \ldots x_n| = |x_1| \cdot |x_2| \ldots |x_n|$.
5. If x and y are real numbers, then show that

   $$\frac{|x+y|}{1+|x+y|} \le \frac{|x|}{1+|x|} + \frac{|y|}{1+|y|}.$$
6. Prove that

   $$|x+y| = |x| + |y| \text{ iff } xy \ge 0$$

   and $|x+y| < |x| + |y|$ iff $xy < 0$.

2 Open Sets, Closed Sets and Countable Sets

1. INTRODUCTION

In this chapter, we shall study the concept of *neighbourhood* of a point, *open and closed sets*, and *limit points* of a set of real numbers and the *Bolzano-Weierstrass theorem*, which is one of the most fundamental theorems of Real Analysis and lays down a sufficient condition for the existence of *limit points* of a set. *We shall be dealing only with real numbers and sets of real numbers unless otherwise stated.*

1.1 Neighbourhood of a Point

A set $N \subseteq \mathbf{R}$ is called the **neighbourhood** of a point a, if there exists an open interval I containing a and contained in N, *i.e.*,

$$a \in I \subseteq N$$

It follows from the definition that an open interval is a neighbourhood of each of its points. Though open intervals containing the point are not the only neighbourhoods of the point but they prove quite adequate for a discussion like ours and are more expressive of the idea of neighbourhoods as understood in ordinary language. We shall, therefore, whenever convenient, take the open interval $]a - \delta, a + \delta[$ where $\delta > 0$ is a neighbourhood of the point a.

Deleted Neighbourhoods

The set $\{x : 0 < |x - a| < \delta\}$, *i.e.*, an open interval $]a - \delta, a + \delta[$ from which the number a itself has been *excluded* or *deleted* is called a *deleted neighbourhood* of a.

Note: For the sake of brevity, we shall write neighbourhood as '**nbd**'.

ILLUSTRATIONS

1. The set $\mathbf{R}$ of real numbers is the neighbourhood of each of its points.
2. The set $\mathbf{Q}$ of rationals is not the *nbd* of any of its points.
3. The open interval $]a, b[$ is *nbd* of each of its points.
4. The closed interval $[a, b]$ is the *nbd* of each point of $]a, b[$ but is not a *nbd* of the end points a and b.

5. The null set ϕ is a *nbd* of each of its points in the sense that there is no point in ϕ of which it is not a *nbd*.

Example 1. A non-empty finite set is not a *nbd* of any point.

A set can be a *nbd* of a point if it contains an open interval containing the point. Since an interval necessarily contains an infinite number of points, therefore, in order that a set be a *nbd* of a point it must necessarily contain an infinity of points. Thus a finite set cannot be a *nbd* of any point.

Example 2. Superset of a *nbd* of a point x is also a *nbd* of x. *i.e.*, if N is a *nbd* of a point x and $M \supseteq N$ then M is also a *nbd* of x.

Example 3. Union (finite or arbitrary) of *nbds* of a point x is again a *nbd* of x.

Example 4. If M and N are *nbds* of a point x, then show that $M \cap N$ is also a *nbd* of x.

■ Since M, N are *nbds* of x, $\exists$ open intervals enclosing the points x such that

$$x \in \left]x - \delta_1, x + \delta_1\right[\subseteq M \text{ and } x \in \left]x - \delta_2, x + \delta_2\right[\subseteq N$$

Let $\delta = \min(\delta_1, \delta_2)$. Then

$$\left]x - \delta, x + \delta\right[\subseteq \left]x - \delta_1, x + \delta_1\right[\subseteq M$$

and

$$\left]x - \delta, x + \delta\right[\subseteq \left]x - \delta_2, x + \delta_2\right[\subseteq N$$

$$\Rightarrow \quad \left]x - \delta, x + \delta\right[\subseteq M \cap N$$

$$\Rightarrow \quad M \cap N \text{ is a } nbd \text{ of } x$$

1.2 Interior Points of a Set

A point x is an *interior point* of a set S if S is a *nbd* of x. In other words, x is an interior point of S if $\exists$ an open interval $]a, b[$ containing x and contained in S, *i.e.*, $x \in \left]a, b\right[\subseteq S$.

Thus a set is a neighbourhood of each of its interior points.

Interior of a Set. The set of all interior points of a set is called the **interior** of the set. The interior of a set S is generally denoted by S^i or int S.

Ex. 1. Show that the interior of the set **N** or **I** or **Q** is the null set, but interior of **R** is **R.**

Ex. 2. Show that the interior of a set S is a subset of S, *i.e.*, $S^i \subseteq S$.

1.3 Open Set

A set S is said to be **open** if it is a *nbd* of each of its points, *i.e.*, for each $x \in S$, there exists an open interval I_x such that

$$mx \in I_x \subseteq S.$$

Thus *every point of an open set is an interior point*, so that for an open set S, $S^i = S$.

Evidently, S is open $\Leftrightarrow S = S^i$

Of course the set is **not open** if it is not a *nbd* of at least one of its points or that there is at least one point of the set which is not an interior point.

ILLUSTRATIONS

1. The set **R** of real numbers is an open set.
2. The set **Q** of rationals is not an open set.
3. The closed interval $[a, b]$ is not open, for it is not a neighbourhood of the end points a and b.
4. The null set ϕ is open, for there is no point in ϕ of which it is not a neighbourhood.
5. A non-empty finite set is not open.
6. The set $\left\{\frac{1}{n}: n \in \mathbf{N}\right\}$ is not open.

Ex. Give an example of an open set which is not an interval.

Example 5. Show that every open interval is an open set. Or, every open interval is a *nbd* of each of its points.

■ Let x be any point of the given open interval $]a, b[$ so that we have $a < x < b$.

Let c, d be two numbers such that

$$a < c < x, \text{ and } x < d < b$$

so that we have

$$a < c < x < d < b \Rightarrow x \in \left]c, d\right[\subset \left]a, b\right[.$$

Thus the given interval $]a, b[$ contains an open interval containing the point x, and is, therefore, a *nbd* of x.

Hence, the open interval is a *nbd* of each of its points and is therefore an open set.

Ex. Show that every point of an open interval is its interior point.

Example 6. Show that every open set is a union of open intervals.

■ Let S be an open set and x_λ a point of S.

Since S is open, therefore $\exists$ an open interval $I_{x\lambda}$ for each of its points x_λ such that

$$x_\lambda \in I_{x_\lambda} \subseteq S \qquad \forall\, x_\lambda \in S$$

Again the set S can be thought of as the union of singleton sets like $\{x_\lambda\}$, *i.e.*,

$$S = \bigcup_{\lambda \in \Lambda} \{x_\lambda\}, \text{ where } \Lambda \text{ is the index set}$$

$$\therefore \qquad S = \bigcup_{\lambda \in \Lambda} \{x_\lambda\} \subseteq \bigcup_{\lambda \in \Lambda} I_{x_\lambda} \subseteq S$$

$$\Rightarrow \qquad S = \bigcup_{\lambda \in \Lambda} I_{x_\lambda}$$

Theorem 1. *The interior of a set is an open set.*

Let S be a given set, and S^i its interior.

If $S^i = \phi$ then S^i is open.

When $S^i \neq \phi$, and let x be any point of S^i.

As x is an interior point of S, $\exists$ an open interval I_x such that $x \in I_x \subseteq S$.

But I_x, being an open interval, is a *nbd* of each of its points.

$\Rightarrow$ every point of I_x is an interior point of I_x, and $I_x \subseteq S$

$\Rightarrow$ every point of I_x is an interior point of S

$\therefore$ $I_x \subseteq S^i$

$\Rightarrow$ $x \in I_x \subseteq S^i \Rightarrow$ any point x of S^i is interior point of S^i

$\Rightarrow$ S^i is an open set.

Corollary. The interior of a set S is an open subset of S.

Theorem 2. *The interior of a set S is the largest open subset of S.*

Or

The interior of a set S contains every open subset of S.

We know that the interior S^i of a set S is an open subset of S. Let us now proceed to show that any open subset S_1 of S is contained in S^i.

Let x be any point of S_1.

Since an open set is a *nbd* of each of its points, therefore S_1 is a *nbd* of x. But S is a superset of S.

$\therefore$ S is also a *nbd* of x

$\Rightarrow$ x is an interior point of S

$\Rightarrow$ $x \in S^i$

Thus, $x \in S_1 \Rightarrow x \in S^i$

$\therefore$ $S_1 \subseteq S^i$

Hence, every open subset of S is contained in its interior S^i.

$\Rightarrow$ S^i, the interior of S, is the largest open subset of S.

Corollary. Interior of a set S is the union of all open subsets of S.

Theorem 3. *The* **union** *of an arbitrary family of open sets is open.*

Let F be the union of an arbitrary family $\mathbf{F} = \{S_\lambda : \lambda \in \Lambda\}$ of open sets, Λ being an index set. To prove that F is open, we shall show that for any point $x \in F$, it contains an open interval containing x.

Let x be any point of F. Since F is the union of the members of $\mathbf{F}$, $\exists$ at least one member, say S_λ of $\mathbf{F}$ which contains x. Again, S_λ being an open set, $\exists$ an open interval I_x such that $x \in I_x \subseteq S_\lambda \subseteq F$.

Thus the set F contains an open interval containing any point x of $F \Rightarrow F$ is an open set.

Theorem 4. *The* **intersection** *of any finite number of open sets is open.*

Let us consider two open sets S and T.

If $S \cap T = \phi$, it is an open set.

If $S \cap T \neq \phi$, let x be any point of $S \cap T$.

Now $x \in S \cap T \Rightarrow x \in S \wedge x \in T$

$\Rightarrow$ S, T are *nbds* of x $[\because S, T \text{ are open}]$

$\Rightarrow$ $S \cap T$ is a *nbd* of x

But since x is any point of $S \cap T$, therefore $S \cap T$ is a *nbd* of each of its points. Hence, $S \cap T$ is open.

The proof may, of course, be extended to a finite number of sets.

Note: The above theorem does not hold for the intersection of *arbitrary* family of open sets.

Consider, for example, the open sets

$$S_n = \left]-\frac{1}{n}, \frac{1}{n}\right[, n \in \mathbf{N}$$

Their intersection is the set $\{0\}$ consisting of the single point 0, and this set is not open.

2. LIMIT POINTS OF A SET

Definition 1. A real number ξ is a **limit point** of a set S $(\subset \mathbf{R})$ if every *nbd* of ξ contains an infinite number of members of S.

Thus ξ is a limit point of a set S if for any *nbd* N of ξ, $N \cap S$ is an infinite set.

A limit point is also called a *cluster point, a condensation point* or an *accumulation point.*

A limit point of a set may or may not be a member of the set. Further, it is clear from the definition that a finite set cannot have a limit point. Also it is not necessary that an infinite set must possess a limit point. In fact a set may have no limit point, a unique limit point, a finite or an infinite number of limit points. A sufficient condition for the existence of a limit point is provided by *Bolzano-Weierstrass theorem* which is discussed in the next section. The following is another definition of a limit point.

Definition 2. A real number ξ is a **limit point** of a set S $(\subseteq \mathbf{R})$ if every *nbd* of ξ contains at least one member of S other than ξ.

The essential idea here is that the points of S different from ξ get 'arbitrarily close' to ξ or 'pile up' at ξ.

Evidently definition 1 implies definition 2. Let us now prove that definition 2 implies definition 1.

Let ξ be a limit point of the set S $(\subseteq \mathbf{R})$ such that every *nbd* of ξ contains at least one point of S other than ξ. Let $]\xi - \delta_1, \xi + \delta_1[$ be one such *nbd* of ξ which contains at least one point, say, $x_1 \neq \xi$ of S.

Let $|x_1 - \xi| = \delta_2 < \delta_1$. Now consider the *nbd* $]\xi - \delta_2, \xi + \delta_2[$ of ξ which by def. 2 of a limit point, must have one point, say, x_2 of S other than ξ.

By repeating the argument with the *nbd* $]\xi - \delta_3, \xi + \delta_3[$ of ξ where $\delta_3 = |x_2 - \xi|$ and so on, it follows that the *nbd* $]\xi - \delta_i, \xi + \delta_i[$ of ξ contains an infinity of members of S.

Hence, Def. 2 $\Rightarrow$ Def. 1.

It is instructive to note that a point ξ is not a limit point of a set S if $\exists$ even one *nbd* of ξ not containing any point of S other than ξ.

Ex. Give a bounded set having (*i*) no limit point, (*ii*) infinite numbers of limit points.

Derived Sets. The set of all limit points of a set S is called the *derived set* of S and is denoted by S'.

ILLUSTRATIONS

1. The set **I** has no limit point, for a *nbd* $\left]m-\frac{1}{2}, m+\frac{1}{2}\right[$ of $m \in \mathbf{I}$, contains no point of **I** other than m. Thus the derived set of **I** is the null set ϕ.
2. Every point of **R** is a limit point, for every *nbd* of any of its points contains an infinite members of **R**. Therefore $\mathbf{R}' = \mathbf{R}$.
3. Every point of the set **Q** of rationals is a limit point, for, between any two rationals there exist infinite rationals. Further every irrational number is also a limit point of **Q** for between any two irrationals there are infinitely many rationals. Thus every real number is a limit point of **Q**, so that $\mathbf{Q}' = \mathbf{R}$.
4. The set $\left\{\frac{1}{n}: n \in \mathbf{N}\right\}$ has only one limit point, zero, which is not a member of the set.
5. Every point of the closed interval $[a, b]$ is its limit point, and a point not belonging to the interval is not a limit point. Thus the derived set $[a, b]' = [a, b]$.
6. Every point of the open interval $]a, b[$ is its limit point. The end points a, b which are not members of $]a, b[$ are also its limit points. Thus
$$]a, b[' = [a, b]$$

Ex. Obtain the derived sets:

1. $\{x: 0 \le x < 1\}$,
2. $\{x: 0 < x < 1, x \in \mathbf{Q}\}$,
3. $\left\{1, -1, 1\frac{1}{2}, -1\frac{1}{2}, -1\frac{1}{3}, \ldots\right\}$,
4. $\left\{1 + \frac{1}{n}: n \in \mathbf{N}\right\}$,
5. $\left\{\frac{1}{m} + \frac{1}{n}: m \in \mathbf{N}, n \in \mathbf{N}\right\}$.

2.1 A finite set has no limit point. Also we have seen that an infinite set may or may not have limit points. We shall now discuss a theorem which sets out sufficient conditions for a set to have limit points.

Bolzano-Weierstrass Theorem (for sets). *Every infinite bounded set has a limit point.*

Let S be any infinite bounded set and m, M its infimum and supremum respectively. Let P be a set of real numbers is defined as follows:

$x \in P$ iff it exceeds at the most a finite number of members of S.

The set P is non-empty, for $m \in P$. Also M is an upper bound of P, for no number greater than or equal to M can belong to P. Thus the set P is non-empty and is bounded above. Therefore, by the order-completeness property, P has the supremum, say ξ. We shall now show that ξ is a limit point of S.

Consider any *nbd*. $]\xi - \varepsilon, \xi + \varepsilon[$ of ξ, where $\varepsilon > 0$.

Since, ξ is the supremum of P, $\exists$ at least one member say η of P such that $\eta > \xi - \varepsilon$. Now η belongs to P, therefore it exceeds at the most a finite number of members of S, and consequently $\xi - \varepsilon\ (< \eta)$ can exceed at the most a finite number of members of S.

Again as ξ is the supremum of P, $\xi + \varepsilon$ cannot belong to P, and consequently $\xi + \varepsilon$ must exceed an infinite number of members of S.

Now, $\xi - \varepsilon$ exceeds at the most a finite number of members of S and $\xi + \varepsilon$ exceeds infinitely many members of S.

$\Rightarrow$ $]\xi - \varepsilon, \xi + \varepsilon[$ contains an infinite number of members of S

Consequently ξ is a limit point of S.

Note: Boundedness is not necessary in order for an infinite set S to have a limit point. The set $S = \left\{\frac{1}{2}, 2, \frac{1}{3}, 3, \frac{1}{4}, 4, \frac{1}{5}, 5 \ldots\right\}$ is unbounded and infinite and has the limit point 0. The unbounded interval $]a, \infty[$ has infinitely many limit points.

2.2 **Example 7.** If S and T are subsets of real numbers, then show that

(*i*) $S \subseteq T \Rightarrow S' \subseteq T'$, and

(*ii*) $(S \cup T)' = S' \cup T'$

■ (*i*) If $S' = \phi$, then evidently $S' \subseteq T'$.

When $S' \neq \phi$, let $\xi \in S'$ and N be any *nbd* of ξ.

$\Rightarrow$ N contains an infinite number of members of S.

But $S \subseteq T$

$\therefore$ N contains infinitely many members of T

$\Rightarrow$ ξ is limit point of T, *i.e.*, $\xi \in T'$.

Thus, $\xi \in S' \Rightarrow \xi \in T'$. Hence, $S' \subseteq T'$.

(*ii*) Now, $S \subseteq S \cup T \Rightarrow S' \subseteq (S \cup T)'$

and $T \subseteq S \cup T \Rightarrow T' \subseteq (S \cup T)'$

Consequently, $S' \cup T' \subseteq (S \cup T)'$ ...(1)

Now we proceed to show that $(S \cup T)' \subseteq S' \cup T'$.

If $(S \cup T)' = \phi$, then evidently $(S \cup T)' \subseteq S' \cup T'$.

When $(S \cup T)' \neq \phi$, let $\xi \in (S \cup T)'$.

Now ξ is a limit point of $(S \cup T)$, therefore, every *nbd* of ξ contains an infinite number of points of $(S \cup T) \Rightarrow$ every *nbd* of ξ contains infinitely many points of S or T or both.

$\Rightarrow$ ξ is a limit point of S or a limit point of T

$\Rightarrow$ $\xi \in S' \vee \xi \in T' \Rightarrow \xi \in S' \cup T'$.

Thus, $\xi \in (S \cup T)' \Rightarrow \xi \in S' \cup T'$

Consequently, $(S \cup T)' \subseteq S' \cup T'$ ...(2)

From (1) and (2), it follows that

$$(S \cup T)' = S' \cup T'$$

Thus the derived set of the union = the union of the derived sets.

Aliter. To show that $(S \cup T)' \subseteq S' \cup T'$

We may show that $\xi \notin S' \cup T' \Rightarrow \xi \notin (S \cup T)'$.

Now $\xi \notin S' \cup T'$ implies that ξ does not belong to either.

$\Rightarrow$ ξ is not a limit point of S or of T

$\therefore \exists$ *nbds* N_1, N_2 of ξ such that N_1 contains no point of S other than ξ and N_2 contains no point of T other than possibly ξ.

Again, since $N_1 \cap N_2 \subseteq N_1$, $N_1 \cap N_2 \subseteq N_2$ therefore $\exists$ a *nbd* $N_1 \cap N_2$ of ξ which contains no point other than ξ of S or of T and thus of $S \cup T$

$\Rightarrow$ ξ is not a limit point $S \cup T$

$\Rightarrow$ $\xi \notin (S \cup T)'$

Thus, $\xi \notin S' \cup T' \Rightarrow \xi \notin (S \cup T)'$

so that $(S \cup T)' \subseteq S' \cup T'$

Example 8. (*i*) If S, T are subsets of $\mathbf{R}$, then show that

$$(S \cap T)' \subseteq S' \cap T'$$

(*ii*) Give an example to show that $(S \cap T)'$ and $S' \cap T'$ may not be equal.

■ (*i*) Now $S \cap T \subseteq S \Rightarrow (S \cap T)' \subseteq S'$ and

$$S \cap T \subseteq T \Rightarrow (S \cap T)' \subseteq T'$$

Consequently, $(S \cap T)' \subseteq S' \cap T'$

(*ii*) Let $S =]1, 2[$ and $T =]2, 3[$, so that

$$S \cap T = \phi \Rightarrow (S \cap T)' = \phi' = \phi.$$

Also $S' = [1, 2], T' = [2, 3]$

$\therefore$ $S' \cap T' = \{2\}$.

Thus, $(S \cap T)' \neq S' \cap T'$.

3. CLOSED SETS: CLOSURE OF A SET

3.1 A real number ξ is said to be an **adherent point** of a set S ($\subseteq \mathbf{R}$) if every *nbd* of ξ contains at least one point of S.

Evidently an adherent point may or may not belong to the set and it may or may not be a limit point of the set.

It follows from the definition that a number $\xi \in S$ is automatically an adherent point of the set, for every *nbd* of a member of the set contains atleast one member of the set, namely the member itself. Further a number $\xi \notin S$ is an adherent point of S only if ξ is a limit point of S, for every *nbd* of ξ, there contains atleast one point of S which is other than ξ.

Thus the set of adherent points of S consists of S and the derived set S'.

The set of all adherent point of S, called the **closure** of S is denoted by $\tilde{S}$, and is such that

$$\tilde{S} = S \cup S'.$$

ILLUSTRATIONS

1. $\tilde{\mathrm{I}} = \mathrm{I} \cup \mathrm{I}' = \mathrm{I} \cup \phi = \mathrm{I}.$
2. $\tilde{\mathrm{Q}} = \mathrm{Q} \cup \mathrm{Q}' = \mathrm{Q} \cup \mathrm{R} = \mathrm{R}.$
3. $\tilde{\mathrm{R}} = \mathrm{R} \cup \mathrm{R}' = \mathrm{R} \cup \mathrm{R} = \mathrm{R}$
4. $\tilde{\phi} = \phi \cup \phi' = \phi \cup \phi = \phi.$

3.2 Closed Sets

A set is said to be **closed** if each of its limit points is a member of the set.

In other words a set S is *closed* if no limit point of S exists which is not contained in S. In rough terms, a set is closed if its points do not get arbitrarily close to any point outside it.

Thus a set S is *closed* iff

$$S' \subseteq S \quad \text{or} \quad \tilde{S} = S.$$

Consequently, a *closed set* is also defined as a set S for which

$$\tilde{S} = S.$$

It should be clearly understood that the concept of closed and open sets are neither mutually exclusive nor exhaustive. The word *not closed* should not be considered equivalent to *open*. Sets exist which are both open and closed, or which are neither open nor closed. The set consisting of points of $]a, b]$ is neither open nor closed.

ILLUSTRATIONS

1. $[a, b]$ is a set which is closed but not open.
2. The set $[0, 1] \cup [2, 3]$, which is not an interval, is closed.
3. The null set ϕ is closed for there exists no limit point of ϕ which is not contained in ϕ. As shown earlier, ϕ is also open.
4. The set $\mathbf{R}$ of real numbers is open as well as closed.
5. The set $\mathbf{Q}$ is not closed, for $\mathbf{Q}' = \mathbf{R} \not\subseteq \mathbf{Q}$. Also it is not open.

6. $\left\{\frac{1}{n} : n \in \mathbf{N}\right\}$ is not closed, for it has one limit point 0, which is not a member of the set. Also it is not open.
7. Every finite set A is a closed set, for its derived set $A' = \phi \subset A$.
8. A set A which has no limit point coincides with its closure, for $A' = \phi$ and $\tilde{A} = A \cup A' = A$.

3.3 Typical Examples

Example 9. Show that the set $S = \{x\colon 0 < x < 1, x \in \mathbf{R}\}$ is open but not closed.

■ The set S is the open interval $]0, 1[$.

$\therefore$ It contains a *nbd* of each of its points. Hence it is an open set.

Again every point of S is a limit point. The end points 0 and 1 which are not members of the set are also limit points. Thus S is not closed.

Example 10. Show that the set

$$S = \left\{1, -\frac{1}{2}, -\frac{1}{2}, \frac{1}{3}, -\frac{1}{3}, \ldots\right\}$$

is neither open nor closed.

■ The members of S heap or cluster near zero on both sides of it and every *nbd* of zero contains an infinite number of points of S. Thus $0 \notin S$ is a limit point $\Rightarrow$ S is not closed.

Again S is not open for it does not contain any *nbd* of any of its points. For example, a *nbd* $\left]\frac{1}{3} - \frac{1}{100}, \frac{1}{3} + \frac{1}{100}\right[$ of $\frac{1}{3}$ is not contained in the set. Hence the set is not open.

Example 11. Show that the set

$$\left\{1, -1, 1\frac{1}{2}, -1\frac{1}{2}, 1\frac{1}{3}, -1\frac{1}{3}, \ldots\right\}$$

is closed but not open.

■ 1 and -1 are the only limit points of the set and are in the set. Therefore, the set is closed.

Again all members of the set (except $1, -1$) are not the interior points of the set. Thus the set is not open.

Hence, the set is closed but not open.

The relationship between closed and open sets is brought out by Theorem 5 that follows and is sometimes taken as the **definition** of a closed set.

3.4 Dense Sets

A subset A of the set of reals $\mathbf{R}$ is said to be *dense* (or *dense* in $\mathbf{R}$ or *everywhere dense*) if every point of $\mathbf{R}$ is a point of A or a limit point of A or equivalently if the closure of A is $\mathbf{R}$.

A set A is said to be *dense in itself* if every point of A is a limit point of A, *i.e.*, if $A \subseteq A'$. A set which is *dense in itself* has no isolated points.

A set A is said to be *nowhere dense* (nondense) relative to $\mathbf{R}$ if no neighbourhood in $\mathbf{R}$ is contained in the closure of A. In other words, the complement of the closure of A is dense in $\mathbf{R}$.

It is clear that if A is an interval or contains an interval then A is not nowhere dense. Because there exists an interval $I \subset \mathbf{R}$ such that $I \cap A \neq \phi$. But there are sets which contain no interval and which fail to be nowhere dense; for example, the set of rationals $\mathbf{Q}$ and the set of irrationals $\mathbf{R} - \mathbf{Q}$.

A set is said to be *perfect* if it is identical with its derived set or equivalently a set which is closed and dense in itself.

ILLUSTRATIONS

1. $A = \left\{1, \frac{1}{2}, \frac{1}{3}, ..., \frac{1}{n}, ...\right\}$, $A' = \{0\}$. The set is neither closed nor dense in itself.
2. The set $\mathbf{Q}$ of rationals is dense in itself but not closed.
3. A finite set is closed but not dense in itself.
4. The set $\mathbf{R}$ of real numbers is dense in itself and closed.
5. The set $\{x: a \leq x \leq b\}$ is a perfect set.
6. The set $\mathbf{R}$ of real numbers is a perfect set.
7. The set $A = \left\{0, 1, \frac{1}{2}, ...\right\}$ is nowhere dense in $\mathbf{R}$ since 0 is the only limit point of A and no neighbourhood of 0 in $\mathbf{R}$ is contained in the closure of A.
8. The set $\mathbf{Q}$ of rational numbers and the set of irrationals are dense in $\mathbf{R}$.
9. The empty set is perfect.

3.5 Some Important Theorems

Theorem 5. *A set is closed iff its complement is open.*

Necessary. Let S be a closed set. We shall show that its complement $\mathbf{R} - S = T$ is open. Let x be *any* point of T.

$$x \in T \Rightarrow x \notin S.$$

Also, since S is closed, x cannot be a limit point of S. Thus $\exists$ a *nbd* N of x such that

$$N \cap S = \phi.$$

$\Rightarrow$ $N \subseteq T \Rightarrow$ every point of T is an interior point.

Thus T is an open set.

Sufficient. Let S be a set whose complement $\mathbf{R} - S = T$ is open.

To show that S is closed, we shall show that every limit point of S is in S.

Let, if possible, a limit point ξ of S be not in S so that ξ is in T. As T is open, $\exists$ a *nbd* of ξ contained in T and thus containing no point of S.

$\therefore$ $\exists$ a *nbd* N of ξ which contains no point of S.

$\Rightarrow$ ξ is not a limit point of S, which is a contradiction.

Hence no limit point of S exists which is not in S.

$\therefore$ S is closed.

Theorem 6. *The intersection of an* **arbitrary** *family of closed sets is closed.*

Let F be the intersection set of an arbitrary family $\mathbf{F} = \{S_\lambda : \lambda \in \Lambda\}$ of closed sets, Λ being an index set.

If the derived set F' of F is ϕ, *i.e.*, when F is a finite set or an infinite set without limit points, then evidently it is closed.

When $F' \neq \phi$, let $\xi \in F'$, *i.e.* ξ be a limit point of F, so that every *nbd* of ξ contains infinitely many members of F and as such of each member S_λ of the family $\mathbf{F}$ of closed sets.

$\Rightarrow$ ξ is limit point of each closed set S_λ

$\Rightarrow$ ξ belongs to each $S_\lambda \Rightarrow \xi \in \bigcap_{\lambda \in \Lambda} S_\lambda = F$.

Thus the set F is closed.

Note: We have given an independent proof but on taking complements, this theorem follows from theorem 3.

Theorem 7. *The union of* **two** *closed sets is a closed set.*

Let S and T be the two given closed sets and ξ a limit point of F, where $F = S \cup T$.

We have to show that $\xi \in F$, for then, the set F will be closed.

Let if possible $\xi \notin F$, thus $\xi \notin S \wedge \xi \notin T$. Also as S and T are closed sets, the point ξ which does not belong to them, cannot be a limit point of either.

$\therefore$ $\exists$ *nbds* N_1 and N_2 of ξ such that

$$N_1 \cap S = \phi \wedge N_2 \cap T = \phi. \quad \text{...(1)}$$

Let $N_1 \cap N_2 = N$, where $\xi \in N$.

$\therefore$ From (1), it follows that

$$N \cap (S \cup T) = \phi \Rightarrow N \cap F = \phi.$$

Thus, $\exists$ a *nbd* N of ξ which contains no point of F.

$\Rightarrow$ ξ is not a limit point of F, which is a contradiction.

Hence, no point not belonging to F can be its limit point, and consequently $F = S \cup T$ is a closed set.

Remarks:

1. The theorem can be extended to the union of a finite number of sets. So we may restate the theorem as: *The union of a finite number of closed sets is closed.*
2. We have given an independent proof but the theorem follows from theorem 4 on taking complements.
3. The union of an arbitrary family of closed sets may not always be a closed set. For example,

let $$S_n = \left[a + \frac{1}{n}, a + 2\right], \text{ for } n \in \mathbf{N} \wedge a \in \mathbf{R}.$$

Then, $\bigcup_{n \in \mathbf{N}} S_n = \left]a, a + 2\right]$, which is not a closed set.

Theorem 8. *The derived set of a set is closed.*

Let S' be the derived set of a set S.

We have to show that the derived set S'' of S' is contained in S'.

Now if $S'' = \phi$, *i.e.*, when S' is either a finite set or an infinite set without limit points, then $S'' = \phi \subset S'$ and therefore S' is closed.

When $S'' \neq \phi$, let $\xi \in S''$, *i.e.*, ξ be a limit point of S'.

$\therefore$ Every *nbd* N of ξ contains at least one point $\eta \neq \xi$ of S'.

Again,

$\eta \in S' \Rightarrow \eta$ is a limit point of S

$\Rightarrow$ every *nbd* of η, N being such a *nbd*, contains infinitely many points of S.

Thus every *nbd* N (of ξ) contains an infinitely many points of S.

$\Rightarrow$ ξ is a limit point of S, *i.e.*, $\xi \in S'$.

Consequently $\xi \in S'' \Rightarrow \xi \in S'$

$\therefore$ $S'' \subseteq S'$, *i.e.*, S' is a closed set.

Corollary 1. S'' is closed, and therefore the closure of S' is S', *i.e.*, $\tilde{S} = S'$.

Corollary 2. For every set S the closure $\tilde{S}$ is closed.

We have simply to show that $(\tilde{S})' \subset \tilde{S}$. Now,

$$(\tilde{S})' = (S \cup S')' = S' \cup S'' = S' \subset \tilde{S}. \qquad \text{(Ref. § 2.2 and Theorem 8)}$$

Theorem 9. *A closed set either contains an interval or else is nowhere dense.*

Let A be any closed set and A is not nowhere dense in $\mathbf{R}$. Then there is some interval I such that for each interval $I_1 \subseteq I$, we have $I_1 \cap A \neq \phi$. We shall show that $I \subseteq A$.

Let $x \in I$. Then every neighbourhood of x contains within it at least one point of A. This implies that either $x \in A$ or else x is a limit point of A. Since A is closed it contains all its limit points and so $x \in A$.

Theorem 10. *The supremum (infimum) of a bounded non-empty set* $\mathbf{S}$ $(\subseteq \mathbf{R})$, *when not a member of* $\mathbf{S}$, *is a limit point of S.*

Let M be the supremum of the bounded set S $(\subseteq \mathbf{R})$, which must exist by the *order completeness* property of $\mathbf{R}$. If $M \notin S$, then for any number $\varepsilon > 0$, however small, $\exists$ at least one member x of S such that

$$M - \varepsilon < x < M.$$

Thus every *nbd* of M contains atleast one member x of the set S other than M. Hence M is a limit point of S.

Corollary. The supremum (infimum) M of a bounded set S is always a member of the closure $\tilde{S}$ of S.

When $M \in S$,

$$M \in S \Rightarrow M \in S \cup S' = \tilde{S} \quad M \in S \Rightarrow M \in S \cup S' = \tilde{S}$$

When $M \notin S$,

$$M \notin S \Rightarrow M \in S' \Rightarrow M \in S \cup S' = \tilde{S}$$

Consequently $M \in \tilde{S}$.

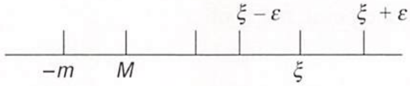
CHAPTER 2

Theorem 11. *The derived set of a bounded set is bounded.*

Let m, M be the bounds of a set S.

It will now be shown that no limit point of S can be outside the interval $[m, M]$.

Let, if possible, $\xi > M$ be a limit point of S, and ε be a positive number such that $\varepsilon < \xi - M$.



Then since M is an upper bound of S, no member of S can lie in the interval $]\xi - \varepsilon, \xi + \varepsilon[$, therefore $\exists$ a *nbd* of ξ which contains no point of S so that ξ cannot be a limit point of S.

Hence, S has no limit point greater than M.

Similarly, it can be shown that no limit point of S is less than m.

Hence, $S' \subseteq [m, M]$.

Corollary. If S is bounded then so is its closure $\tilde{S}$.

$S \subseteq [m, M] \Rightarrow S' \subseteq [m, M] \Rightarrow \tilde{S} = S \cup S' \subseteq [m, M]$.

Remark: If supremum M (infimum m) of S is not a member of S, then it is a limit point of S and in view of the above theorem, it is the greatest (least) member of S'.

However, if it is a member of S, then it is not necessarily a limit point of S, so that M (or m) may not be a member of $S' \subseteq [m, M]$. Thus M, m may not always be supremum and infimum of S' but they are always so for $\tilde{S} = S \cup S'$.

For example, for the set $S = \left\{-1, 1, -1\frac{1}{2}, 1\frac{1}{2}, -1\frac{1}{3}, 1\frac{1}{3}, ...\right\}$

$$m = -1\frac{1}{2},\ M = 1\frac{1}{2},\ S' = \{-1, 1\}$$

$$\inf S' = -1 \neq m,\ \sup S' = 1 \neq M$$

but

$$\inf \tilde{S} = m, \sup \tilde{S} = M.$$

Theorem 12. *The derived set S' of a bounded infinite set S ($\subseteq \mathbf{R}$) has the smallest and the greatest members.*

Since the set S is bounded, therefore S' is also bounded. Also S' is non-empty, by Bolzano-Weierstrass theorem S has at least one limit point.

Now S' may be finite or infinite.

When $S'(\neq \phi)$ is finite, evidently it has the greatest and the least members.

When S' is infinite, being bounded set of real numbers, by order-completeness property of $\mathbf{R}$, it has the supremum G and the infimum g.

It will now be shown that G, g are limit points of S, *i.e.*,

$$G \in S', \quad g \in S'$$

Let us first consider G.

Let $]G - \varepsilon, G + \varepsilon[, \varepsilon > 0$ be any *nbd* of G.

Now G being the supremum of S', $\exists$ at least one member ξ of S' such that $G - \varepsilon < \xi \leq G$.

Thus $]G-\varepsilon, G+\varepsilon[$ is a *nbd* of ξ.

But ξ is a limit point of S, so that $]G-\varepsilon, G+\varepsilon[$ contains an infinite members of S.

$\Rightarrow$ any *nbd* $]G-\varepsilon, G+\varepsilon[$ of G contains an infinite number of members of S.

$\Rightarrow$ G is a limit point of $S \Rightarrow G \in S'$

Similarly, it can be shown that $g \in S'$.

Thus, $G \in S'$ and $g \in S'$, being supremum and infimum of S', are the greatest and the smallest members of S'.

The theorem may be restated as:

Every bounded infinite set has the smallest and the greatest limit points.

The smallest and greatest limit points of a set are called the **lower** and **upper limits of indetermination** or simply the **lower** and **upper limits** respectively of the set.

4. COUNTABLE AND UNCOUNTABLE SETS

An infinite set A is said to be *Countably infinite* (or denumerable or enumerable) if it is equivalent to the set **N** of natural numbers.

A set which is either empty, finite or countably infinite is called *countable* otherwise it is *uncountable*.

ILLUSTRATIONS

1. The set of all integers is countable.
2. The set $\{1, 4, 9, 16, ...\}$ is countable.
3. The set P_n of all polynomial functions with integer coefficients is countable.
4. The set of all ordered pairs of integers is countable.
5. The set of all real numbers is uncountable.

Example 12. The set of real numbers in $[0, 1]$ is uncountable.

■ Let the set of all real numbers in $[0, 1]$ be countable, *i.e.*, $\{x : 0 \le x \le 1\} = \{x_1, x_2, ...\ x_n, ...\}$. Each real number x_i in $[0, 1]$ has a decimal expansion $0, a_1, a_2, ...\ a_n, ...$where a_i, $i \in \mathbf{N}$, are any of the digits 0, 1, 2, ..., 9. We assume that the numbers whose decimal expansion terminate such as .0573 are written as .0573000 ... which is the same as .0572999.... Since all real numbers in $[0, 1]$ are countable, therefore, we can establish a 1–1 correspondence of the members of $[0, 1]$ with the set of positive integers in the following manner:

$$1 \leftrightarrow 0.a_{11}a_{12}a_{13}...$$

$$2 \leftrightarrow 0.a_{21}a_{22}a_{23}...$$

$$3 \leftrightarrow 0.a_{31}a_{32}a_{33}...$$

....

We now construct a number $0.b_1b_2b_3...$ where $** \ b_i = \begin{cases} 4 \text{ if } a_{ii} = 5 \\ \qquad i = 1, 2, 3, ... \\ 5 \text{ if } a_{ii} \ne 5 \end{cases}$

(any two digits can be used instead of 4 and 5). Then the number $0.b_1b_2b_3...$ lies between 0 and 1 and is different from the numbers in the above list and therefore cannot be in the list, contradicting the assumption that the set of all real numbers in [0, 1] is countable.

Example 13. The set of rational numbers in [0, 1] is countable.

■ In order to show that the set of rational numbers in [0, 1] is countable, we must show that there exists a 1–1 correspondence between the set of rationals of [0, 1] and the set of positive integers. Arrange the set of rationals according to increasing denominators as

$0, 1, \frac{1}{2}, \frac{1}{3}, \frac{2}{3}, \frac{1}{4}, \frac{3}{4}, \frac{1}{5}, \frac{2}{5}, \frac{3}{5}, \frac{4}{5}, \frac{1}{6}, \frac{5}{6}, ...$, etc. Then the 1–1 correspondence can be indicated as:

$1 \leftrightarrow 0$	$5 \leftrightarrow \frac{2}{3}$	$9 \leftrightarrow \frac{2}{5}$
$2 \leftrightarrow 1$	$6 \leftrightarrow \frac{1}{4}$	$10 \leftrightarrow \frac{3}{5}$
$3 \leftrightarrow \frac{1}{2}$	$7 \leftrightarrow \frac{3}{4}$	$11 \leftrightarrow \frac{4}{5}$
$4 \leftrightarrow \frac{1}{3}$	$8 \leftrightarrow \frac{1}{5}$	....

Theorem 13. If ** $f: A \to B$ *is one-to-one one B is countable then A is countable.*

If A is finite, then there is nothing to prove. Suppose A is infinite. Now A is equivalent to $f(A)$ where $f(A)$ is the range of f, so $f(A)$ is infinite. Also $f(A) \subseteq B$. Therefore B is infinite. By hypothesis B is countable so B is countably infinite. Define a mapping $\phi: \mathbf{N} \to B$ by $\phi(n) = b_n$ for each $n \in \mathbf{N}$. Then $B = \{b_1, b_2, ...\}$. Let n_1 be the first natural number such that $b_{n_1} \in f(A)$. Let n_2 be the first natural number greater than n_1 such that $b_{n_2} \in f(A)$. Again n_3 be the first natural number greater than n_2 such that $b_{n_3} \in f(A)$ and so on.

Thus $f(A) = \{b_{n_1}, b_{n_2}, b_{n_3}, ... b_{n_k}, ...\}$. We now define a mapping $g: f(A) \to \mathbf{N}$ by

$g(b_{n_k}) = k$, for $k = 1, 2, 3, ...$

The mapping 'g' is one-to-one and onto for $k \neq j, n_k \neq n_j$. Also $f: A \to B$ and $g: f(A) \to \mathbf{N}$ implies the composition mapping $gof: A \to \mathbf{N}$ is one-one and onto. Thus A is equivalent to $\mathbf{N}$ and hence countable.

Corollary. Every subset of a countable set is countable.

Theorem 14. *The cartesian product of two countable sets is countable.*

Let A and B be any two countable sets. Then $A \times B = \{(a, b): a \in A, b \in B\}$ is their cartesian product. Now if any of the two sets is empty then $A \times B = \phi$ and there is nothing to prove. If one of the sets is finite, say A is finite with m elements then the product of

$$A = \{a_1, a_2, ..., a_m\} \text{ and } B = \{b_1, b_2, ..., b_n, ...\}$$

is

$$\begin{aligned} A \times B = \{&(a_1, b_1), (a_1, b_2), ..., (a_1, b_n), ... \\ &(a_2, b_1), (a_2, b_2), ..., (a_2, b_n), ... \\ &\quad\vdots \qquad\quad \vdots \\ &(a_m, b_1), (a_m, b_2), ..., (a_m, b_n), ...\} \end{aligned}$$

which can be seen to be equivalent to **N** by listing the elements as

$(a_1, b_1), (a_2, b_1) \ldots (a_m, b_1); (a_1, b_2), (a_2, b_2) \ldots (a_m, b_2); \ \ldots; (a_1, b_n)\ (a_2, b_n) \ldots (a_m, b_n); \ldots$

Let A and B be both countably infinite

$$A = \{a_1, a_2, \ldots\}$$
$$B = \{b_1, b_2, \ldots\},$$

then $A \times B$ is equivalent to **N** which can be exhibited as

$(a_1, b_1); (a_2, b_1), (a_1, b_2); (a_3, b_1), (a_2, b_2), (a_1, b_3); (a_4, b_1), (a_3, b_2), (a_2, b_3); (a_1, b_4), \ldots$

The function $f : A \times B \to \mathbf{N}$ is defined as

$$f(a_1, b_1) = 1, f(a_2, b_1) = 2, f(a_1, b_2) = 3 \text{ and so on,}$$

f is one-to-one and onto. Therefore $A \times B$ is countably infinite.

Theorem 15. *A countable union of countable sets is countable.*

Consider the sets $A_i = \{a_{1i}, a_{2i}, a_{3i}, \ldots\}$, $i = 1, 2, 3, \ldots$ Each A_i, $i = 1, 2, 3, \ldots$ is countable. The k th element of A_i is a_{ki}. The elements of the countable union $\bigcup_{i=1}^{\infty} A_i$ of the sets A_i's can be listed as $a_{11}, a_{12}, a_{21}, a_{13}, a_{22}, a_{31}, a_{14}, a_{23}, a_{32}, a_{41}, \ldots$ (the order has been taken according to the sum $i + j = k$, $k = 2, 3, \ldots$ i, j being the suffices of the element $a_{ij} \in A_j$). The one-one correspondence between the elements of $\bigcup_{i=1}^{\infty} A_i$ and the set of positive integers is given by

$$\begin{array}{cccccccccc} a_{11} & a_{12} & a_{21} & a_{13} & a_{22} & a_{31} & a_{14} & a_{23} & a_{32} & a_{41} \ldots \\ \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \ldots \end{array}$$

Hence, the set $\bigcup_{i=1}^{\infty} A_i$ is countable.

Corollary. The set of all rational numbers is countable.

The set of all rational numbers is the union $\bigcup_{i=1}^{\infty} A_i$, where A_i is the set of rationals which can be written with denominator i. That is $A_i = \left\{\frac{0}{i}, \frac{-1}{i}, \frac{1}{i}, \frac{-2}{i}, \frac{2}{i}, \ldots\right\}$. Each A_i is equivalent to the set of all positive integers and hence countable.

Example 14. The set **R** of real numbers is uncountable.

■ Suppose the set **R** is countable. Then $\mathbf{R} = \{x_1, x_2, x_3, \ldots\}$. Enclose each member x_n of **R** in an open interval $I_n = \left]x_n - \frac{1}{2^{n+1}}, x_n + \frac{1}{2^{n+1}}\right[$ of length $\frac{1}{2^n}$, $n = 1, 2, 3, \ldots$. The sum of the lengths of I_n's is $\frac{1}{2} + \frac{1}{2^2} + \frac{1}{2^3} + \ldots$. But $x_n \in \mathbf{R}$ and $\mathbf{R} = \bigcup_{n=1}^{\infty} \{x_n\} \subseteq \bigcup_{n=1}^{\infty} I_n$. This implies that the whole real line (whose length is infinite) is contained in the union of intervals whose lengths add up to 1. This is a contradiction. Hence the set **R** is uncountable.

3 Real Sequences

In this chapter we shall study a special class of functions whose domain is the set **N** of natural numbers and range a set of real numbers—the *Real Sequences*.

1. SEQUENCES

A function whose domain is the set **N** of natural numbers and range, a set of real numbers is called a *real sequence*. Thus a real sequence is denoted symbolically as $S: \mathbf{N} \to \mathbf{R}$.

Since we shall be dealing with real sequences only, we shall use the term *sequence* to denote a *Real sequence*.

NOTATION: Since the domain for a sequence is always **N**, a sequence is specified by the values S_n, $n \in \mathbf{N}$. Thus a sequence may be denoted as

$$\{S_n\},\ n \in \mathbf{N} \text{ or } \{S_1, S_2, S_3, \ldots S_n, \ldots\}$$

The values $S_1, S_2, S_3, \ldots$ are called the first, second, third ... terms of the sequence. The mth and nth terms S_m and S_n for $m \neq n$ are treated as distinct terms even if $S_m = S_n$. Thus the terms of a sequence are arranged in a definite order as first, second, third, ... terms and the terms occurring at different positions are treated as distinct terms even if they have the same value. The number of terms in a sequence is always infinite.

In other words, we define a sequence as an ordered set of real numbers whose members can be put in an one-one correspondence with the set of natural numbers. However, a sequence may have only a finite number of distinct elements.

For example:

1. $\{S_n\} = \{(-1)^n\}$, $n \in \mathbf{N}$.

   Here $S_1 = -1$, $S_2 = 1$, $S_3 = -1$, $S_4 = 1$, ... so that there are only two, 1, –1 distinct elements.

2. $\{S_n\} = \left\{\dfrac{1}{n}\right\}$, $n \in \mathbf{N}$.

   Here $S_1 = 1,\ S_2 = \dfrac{1}{2}, S_3 = \dfrac{1}{3}, \ldots$

   All the elements are distinct.

ILLUSTRATIONS

1. $\{S_n\}$, where $S_n = \left(1 + \frac{1}{n}\right)^n$, $n \in \mathbf{N}$.
2. $\{S_n\}$, where $S_n = 1 + (-1)^n$, $n \in \mathbf{N}$.
3. $\{S_n\}$, where $S_n = 1, \forall\, n \in \mathbf{N}$.
4. $\left\{\frac{(-1)^{n-1}}{n!}\right\}$, $n \in \mathbf{N}$.

1.1 The Range

The Range or the Range Set is the set consisting of all distinct elements of a sequence, without repetition and without regard to the position of a term. Thus the range may be a finite or an infinite set, without ever being the null set.

1.2 Bounds of a Sequence

Bounded above sequences

A sequence $\{S_n\}$ is said to be *bounded above* if there exists a real number K such that

$$S_n \le K \qquad \forall\, n \in \mathbf{N}$$

Bounded below sequences

A sequence $\{S_n\}$ is said to be *bounded below* if there exists a real number k such that

$$S_n \ge k \quad \forall\, n \in \mathbf{N}.$$

Bounded sequences

A sequence is said to be *bounded* when it is bounded both above and below. K and k are respectively the upper and the lower bounds of the sequence.

Evidently a sequence is bounded iff its range is bounded. Also the bounds of the range are the bounds of the sequence.

1.3 Convergence of Sequences

Definition 1. A sequence $\{S_n\}$ is said to converge to a real number l (or to have the real number l as its limit) if for each $\varepsilon > 0$, there exists a positive integer m (depending on ε) such that $|S_n - l| < \varepsilon$, for all $n \ge m$.

The fact is expressed by saying that the terms approach the value l or tend to l as n becomes larger and larger. The same thing expressed in symbols is

$$S_n \to l \text{ as } n \to \infty \text{ or } \lim_{n\to\infty} S_n = l.$$

The definition ensures that

(*i*) From some stage onwards the difference between S_n and l can be made less than any preassigned positive number ε, however small, *i.e.*, given any positive real number ε, no matter

CHAPTER 3

however small, $\exists$ a positive integer m (finite) such that mth term onwards, S_n becomes and remains arbitrarily close to l, *i.e.*, l is a limit point of the sequence.

(*ii*) For any $\varepsilon > 0$, at the most a finite number of terms (depending on the choice of ε) of the sequence can lie outside $]l - \varepsilon, l + \varepsilon[$, *i.e.*, there is at the most a finite number of n's for which

$$S_n \le l - \varepsilon, \text{ and } S_n \ge l + \varepsilon.$$

(*iii*) Since $l - \varepsilon < S_n < l + \varepsilon$ for all $n \ge m$, therefore $S_n < l + \varepsilon$, for infinite number of terms, *i.e.*, infinite number of terms lie to the left of $l + \varepsilon$, or to the right of $l - \varepsilon$.

It may, therefore, be observed that if we can find even one $\varepsilon > 0$ for which infinitely many terms of the sequence lie outside $]l - \varepsilon, l + \varepsilon[$, then the sequence cannot converge to l.

1.4 Some Theorems

Theorem 1. *Every convergent sequence is bounded.*

Let a sequence $\{S_n\}$ converge to the limit l.

Let $\varepsilon > 0$ be a given number, so that $\exists$ a positive integer m such that

$$|S_n - l| < \varepsilon \quad \forall\, n \ge m$$

$$\Leftrightarrow \qquad l - \varepsilon < S_n < l + \varepsilon \quad \forall\, n \ge m.$$

Let $\quad g = \min\left\{l - \varepsilon, S_1, S_2, \ldots, S_{m-1}\right\}.$

$\quad G = \max\left\{l + \varepsilon, S_1, S_2, \ldots, S_{m-1}\right\}.$

Thus, we have

$$g \le S_n \le G \quad \forall\, n.$$

Hence, $\{S_n\}$ is a bounded sequence.

Remark: The converse of the above theorem may not be true. For example the sequence $\{S_n\}$, where $S_n = (-1)^n$, $n \in \mathbf{N}$, is bounded but it is not convergent. For, if possible, $\lim\limits_{n\to\infty} S_n = l$ then for $\varepsilon = 1, \exists\, m \in \mathbf{N}$ such that

$$\left|S_n - l\right| < 1, \quad \forall\, n \ge m,$$

i.e., $\quad \left|(-1)^{2m} - l\right| < 1 \text{ and } \left|(-1)^{2m+1} - l\right| < 1$

or $\quad \left|1 - l\right| < 1 \text{ and } \left|1 + l\right| < 1$

$\Rightarrow \quad 2 = \left|1 - l + 1 + l\right| < 1 + 1 = 2$

which is absurd.

Theorem 2. *A sequence cannot converge to more than one limit.*

Let, if possible, a sequence $\{S_n\}$ converges to two real numbers l and $l'\,(l \ne l')$. Let us select $\varepsilon = \dfrac{1}{3}|l - l'| > 0$

Since the sequence $\{S_n\}$ converges to l and l'; therefore, there exist positive integers m_1 and m_2 such that

$$\left|S_n - l\right| < \varepsilon, \quad \forall\, n \ge m_1 \qquad \ldots(1)$$

and $$|S_n - l'| < \varepsilon, \quad \forall\, n \geq m_2 \qquad \text{...(2)}$$

Now from (1) and (2), for $n \geq \max(m_1, m_2)$

$$|l - l'| = |l - S_n + S_n - l'| \leq |l - s_n| + |S_n - l'| < 2\varepsilon$$

i.e., $|l - l'| < \frac{2}{3}|l - l'|$, which is not possible.

Hence, the sequence cannot converge to two limits.

It may be seen from the definition that the number to which a sequence converges is a limit point of the sequence. Consequently, the unique limit to which the sequence converges is called *the limit point* or the **limit** of the sequence. Symbolically, we write

$$\lim_{n\to\infty} S_n = l \text{ or } S_n \to l \text{ as } n \to \infty,$$

or simply

$$\lim S_n = l.$$

Thus, *Theorems* 1 *and* 2 may be stated in a combined form as:

Theorem 3. *Every convergent sequence is bounded and has a unique limit.*

2. LIMIT POINTS OF A SEQUENCE

A real number ξ is said to be a **limit point** of a sequence $\{S_n\}$, if every neighbourhood of ξ contains an infinite number of members of the sequence.

Thus ξ is a limit point of a sequence if given any positive number ε, however small, $S_n \in]\xi - \varepsilon, \xi + \varepsilon[$ for an infinite number of values of n, *i.e.*,

$$|S_n - \xi| < \varepsilon, \text{ for infinitely many values of } n.$$

Intuitively, it means that S_n is arbitrarily close to ξ for an infinite number of values of n or that infinitely many terms of the sequence occur very close to ξ.

As in a set, a *limit point* is also called a *cluster point* or a *point of condensation*.

Thus a number ξ is *not a limit point* of the sequence $\{S_n\}$ if $\exists$ a number $\varepsilon > 0$ such that $S_n \in]\xi - \varepsilon, \xi + \varepsilon[$ for at the most a finite number of values of n.

Note: *A* more intuitive but rigorous way of finding a limit point l is to see if the terms get 'closer and closer' to l. This will provide a 'guess' as to the limit point, after which the definition is applied to see if the guess is correct.

It is clear from the definition that a limit point of the range set of a sequence is also a limit point of the sequence. But the converse may not always be true. It may happen that the limit point ξ of a sequence is such that $S_n = \xi$ for an infinitely many values of n, so that automatically $S_n \in]\xi - \varepsilon, \xi + \varepsilon[$ for an infinity of values of n. In such a situation ξ is just one element in the range set and as such fails to be a limit point thereof. However, if $S_n \in]\xi - \varepsilon, \xi + \varepsilon[$ for an infinite number of values of n and $S_n = \xi$ for at the most a finite number of values of n, then a limit point ξ of the sequence is as well a limit point of the range.

ILLUSTRATIONS

1. The constant sequence $\{S_n\}$, where $S_n = 1, \ \forall\ n \in \mathbf{N}$, has the only limit point 1. The range is $\{1\}$ and has no limit point.
2. The sequence $\{S_n\}$, where $S_n = \dfrac{1}{n}, n \in \mathbf{N}$, has 0 as a limit point which is as well a limit point of the range $\left\{1, \dfrac{1}{2}, \dfrac{1}{3}, \dfrac{1}{4}, \ldots\right\}$
3. 0 and 2 are the only limit points of the sequence $\{S_n\}$, where $S_n = 1 + (-1)^n,\ n \in \mathbf{N}$. The range set $\{0, 2\}$ has no limit point.
4. 1 and -1 are the two limit points of the sequence $\{S_n\}$, where $S_n = (-1)^n,\ n \in \mathbf{N}$. The range set $\{1, -1\}$ has no limit point.
5. 1 and -1 are the two limit points of the sequence $\{S_n\}$, where $S_n = (-1)^n \left(1 + \dfrac{1}{n}\right), n \in \mathbf{N}$, which are also the limit points of the range.

2.1 Existence of Limit Points

Since members of a sequence form a set (the range), all the theorems relating to bounds and limit points of sets also hold for sequences of members with suitable modifications. It is not necessary, therefore, to list these again. However, we give a proof of one of the most fundamental theorems for sequences.

Bolzano-Weierstrass Theorem (for sequences)

Every bounded sequence has a limit point.

Let $\{S_n\}$ be a bounded sequence and $S = \{S_n : n \in \mathbf{N}\}$ be its range. Since the sequence is bounded, therefore its range set S is also bounded.

There are two possibilities:

(*i*) S is finite, (*ii*) S is infinite.

(*i*) If S is finite, then there must exist at least one member $\xi \in S$ such that $S_n = \xi$ for an infinite number of values of n. This means that every neighbourhood $]\xi - \varepsilon, \xi + \varepsilon[$ of ξ, contains $S_n\ (= \xi)$ for an infinite number of values of n.

Thus ξ is a limit point of the sequence $\{S_n\}$.

(*ii*) When S is infinite, since it is bounded, it has by Bolzano-Weierstrass theorem (for sets), at least one limit point, say ζ.

Again, since ζ is a limit point of S, therefore every neighbourhood $]\xi - \varepsilon, \xi + \varepsilon[, \varepsilon > 0$ of ζ contains an infinity of members of S, *i.e.*, $S_n \in]\zeta - \varepsilon, \zeta + \varepsilon[$ for an infinity of values of n. Hence ζ is a limit point of the sequence.

Note: The converse of the theorem is not always true, for there do exist unbounded sequences having only one real limit point.

For example $\{1, 2, 1, 4, 1, 6, \ldots\}$ has a unique limit point 1, but is not bounded above.

2.2 We have seen that a bounded sequence has at least one limit point. It may have one limit point, a finite or an infinite number of limit points. A number of questions arise. It may be asked, 'what is the

greatest or the least limit point or whether the greatest or the least limit point exists at all?' In an attempt to answer such questions we now proceed to prove that such points do exist for bounded sequences.

Theorem 4. *The set of the limit points of a bounded sequence has the greatest and the least members.*

Let $\{S_n\}$ be a bounded sequence and S its range set. The set S is bounded and consequently its derived set S' is also bounded (Theorem 11, Ch. 2).

Let T be the set of limit points of the sequence. T is non-empty, for, by Bolzano-Weierstrass theorem the sequence has at least one limit point. Again, since T consists of the limit points of S (*i.e.*, the derived set S') and *those points of* S which are not the limit points of S but are limit points of the sequence. Therefore, T is bounded.

T may, however, be finite or infinite.

When $T(\neq \phi)$ is finite, it evidently has the greatest and the least members.

When T is infinite, being bounded set of real numbers, by the order-completeness property of real numbers, it has the Supremum M and the Infimum m, say.

It will now be shown that M and m are the limit points of the sequence, *i.e.*, $M \in T$, $m \in T$.

Let us first consider M.

Let $]M - \varepsilon, M + \varepsilon[$, $\varepsilon > 0$ be any *nbd* of M.

Since M is the supremum of T, therefore, $\exists$ at least one member ξ of T such that $M - \varepsilon < \xi \leq M$.

Thus $]M - \varepsilon, M + \varepsilon[$ is a *nbd* of ξ.

But ξ is a limit point of the sequence, so that $]M - \varepsilon, M + \varepsilon[$ contains an infinite members of the sequence.

$\Rightarrow$ and *nbd* $]M - \varepsilon, M + \varepsilon[$ of M contains an infinite number of members of the sequence

$\Rightarrow$ M is a limit point of the sequence

$\Rightarrow$ $M \in T$

Similarly it may be shown that $M \in T$.

Thus $M \in T$, $m \in T$ and being the supremum and infimum of T are the greatest and the least members of T respectively.

Thus *a bounded sequence has the greatest and the least limit points* (one farthest to the right and other farthest to the left).

2.3 The greatest and the smallest of the limit points of a (bounded) sequence are respectively called the **upper limit** and the **lower limit**.

ILLUSTRATIONS

1. The sequence $\{S_n\}$, where $S_n = (-1)^n$, $n \in \mathbf{N}$, is bounded, for $-1 \leq S_n \leq 1$, $\forall\, n \in \mathbf{N}$.

   Also -1, 1 are the only limit points.

   $\therefore$ Upper limit $= 1$, lower limit $= -1$.

2. The sequence $\{S_n\}$, where $S_n = 1 + (-1)^n$, $n \in \mathbf{N}$, is bounded, for $0 \leq S_n \leq 2$, $\forall\, n \in \mathbf{N}$.

   0, 2 are the only limit points.

   $\therefore$ Upper limit $= 2$, lower limit $= 0$.

3. The sequence $\{S_n\}$, where $S_n = \dfrac{(-1)^{n-1}}{n!}$, $n \in \mathbf{N}$, *i.e.*, $\left\{1, \dfrac{-1}{2!}, \dfrac{-1}{3!}, \dfrac{-1}{4!}, \ldots\right\}$ is bounded, for $-\frac{1}{2} \le S_n \le 1, \forall\, n$; 0 being the only limit point, the upper and the lower limits coincide with 0, and so $\lim\limits_{n \to \infty} S_n = 0$.
4. The sequence $\{S_n\}$, where $S_n = n^2$, is $(1, 4, 9, 16, 25, \ldots\}$. The sequence is bounded below but not above. There is no real limit point.

3. LIMITS — INFERIOR AND SUPERIOR

From the definition of limit in Section 1.4, it follows that the limiting behaviour of any sequence $\{a_n\}$ of real numbers, depends only on sets of the form $\{a_n : n \ge m\}$, *i.e.*, $\{a_m, a_{m+1}, a_{m+2}, \ldots\}$. In this regard we make the following definition.

Definition. Let $\{a_n\}$ be a sequence of real numbers (not necessarily bounded). We define

$$\lim_{n\to\infty} \inf a_n = \sup_n \inf \left\{a_n, a_{n+1}, a_{n+2}, \ldots\right\}$$

and

$$\lim_{n\to\infty} \sup a_n = \inf_n \sup \left\{a_n, a_{n+1}, a_{n+2}, \ldots\right\}$$

as the limit inferior and limit superior respectively of the sequence $\{a_n\}$.

We shall denote limit inferior and limit superior of $\{a_n\}$ by $\underline{\lim}_{n\to\infty} a_n$ and $\overline{\lim}_{n\to\infty} a_n$ or simply by $\underline{\lim}\, a_n$ and $\overline{\lim}\, a_n$ respectively.

We shall use the following notations for the sequence $\{a_n\}$, for each $n \in \mathbf{N}$.

$$\underline{A}_n = \inf \left\{a_n, a_{n+1}, a_{n+2}, \ldots\right\},$$

and

$$\overline{A}_n = \sup \left\{a_n, a_{n+1}, a_{n+2}, \ldots\right\}.$$

Therefore, we have

$$\underline{\lim}\, a_n = \sup_n \underline{A}_n$$

and

$$\overline{\lim}\, a_n = \inf_n A_n$$

Now $\{a_{n+1}, a_{n+2}, \ldots\} \subseteq \{a_n, a_{n+1}, a_{n+2}, \ldots\}$ therefore by taking infimum and supremum respectively, it follows that

$$\underline{A}_{n+1} \ge \underline{A}_n \text{ and } \overline{A}_{n+1} \le \overline{A}_n$$

This is true for each $n \in \mathbf{N}$.

The above inequalities show that the associated sequences $\{\underline{A}_n\}$ and $\{\overline{A}_n\}$ monotonically increase and decrease respectively with n.

Remark: It should be noted that both limits, inferior and superior, exist uniquely (finite or infinite) for all real sequences.

Theorem 5. *If $\{a_n\}$ is any sequence, then*

$$\inf a_n \le \underline{\lim}\, a_n \le \overline{\lim}\, a_n \le \sup a_n.$$

Let
$$\underline{A}_k = \inf_{n \ge k} a_n \text{ and } \overline{A}_k = \sup_{n \ge k} a_n,\ k \in \mathbf{N}.$$

Then, for all $k,\ n \in \mathbf{N}$, we have

$$\underline{A}_k \le \underline{A}_{k+n} \le \overline{A}_{k+n} \le \overline{A}_n$$

$$\therefore \quad \underline{A}_k \le \overline{A}_n, \text{ for all } k, n \in \mathbf{N}$$

This implies that each $\underline{A}_k$ is a lower bound of the sequence $\{\overline{A}_n\}$, therefore

$$\underline{A}_k \le \inf_n \overline{A}_n = \overline{\lim}\, a_n, \text{ for each } k \in \mathbf{N}$$

This gives $\overline{\lim}\, a_n$ is an upper-bound of the sequence $\{\underline{A}_k\}$. Hence,

$$\underline{\lim}\, a_n = \sup_k \underline{A}_k \le \overline{\lim}\, a_n.$$

Other inequalities follow from

$$\inf a_n = \underline{A}_1 \le \underline{A}_k \le \overline{A}_n \le \overline{A}_1 = \sup a_n$$

and the definition of $\underline{\lim}\, a_n$ and $\overline{\lim}\, a_n$ respectively.

Corollary. *If* a sequence $\{a_n\}$ is bounded then limit inferior and limit superior of $\{a_n\}$ are both finite. In fact

$$-\infty < \underline{\lim}\, a_n \le \overline{\lim}\, a_n < +\infty.$$

Theorem 6. *If $\{a_n\}$ is any sequence, then*

$$\underline{\lim}\,(-a_n) = -\overline{\lim}\, a_n, \text{ and } \overline{\lim}\,(-a_n) = -\underline{\lim}\, a_n.$$

Let $b_n = -a_n, n \in \mathbf{N}$, then we have

$$\begin{aligned} \underline{B}_n &= \inf \{b_n, b_{n+1}, \ldots\} \\ &= -\sup \{a_n, a_{n+1}, \ldots\} = -\overline{A}_n \end{aligned}$$

and so,

$$\begin{aligned} \underline{\lim}\,(-a_n) &= \underline{\lim}\, b_n = \sup\,(\underline{B}_1, \underline{B}_2, \ldots) \\ &= \sup\{-\overline{A}_1, -\overline{A}_2, \ldots\} \\ &= -\inf\{\overline{A}_1, \overline{A}_2, \ldots\} \\ &= -\inf \overline{A}_n = -\overline{\lim}\, a_n. \end{aligned}$$

CHAPTER 3

Also,

$$\underline{\lim}\, a_n = \underline{\lim}\,(-(-a_n)) = -\overline{\lim}\,(-a_n).$$

ILLUSTRATIONS

1. If $a_n = (-1)^n, n \in \mathbf{N}$, then

$$\underline{A}_n = -1 \text{ and } \overline{A}_n = 1, \text{ for each } n \in \mathbf{N}$$

$$\underline{\lim}\, a_n = \sup \underline{A}_n = -1 \text{ and } \overline{\lim}\, a_n = \inf \overline{A}_n = 1.$$

2. If $a_n = 1 + (-1)^n, n \in \mathbf{N}$, then

$$\underline{A}_n = \inf \left\{1 + (-1)^n, 1 + (-1)^{n+1}, \ldots\right\} = 0$$

and $$\overline{A}_n = \sup \left\{1 + (-1)^n, 1 + (-1)^{n+1}, \ldots\right\} = 2, \text{ for each } n \in \mathbf{N}$$

$$\underline{\lim}\, a_n = 0 \text{ and } \overline{\lim}\, a_n = 2.$$

3. If $a_n = n, n \in \mathbf{N}$, then

$$\underline{A}_n = n, \text{ and } \overline{A}_n = +\infty$$

$$\underline{\lim}\, a_n = \sup \{1, 2, 3, \ldots\} = +\infty, \text{ and } \overline{\lim}\, a_n = +\infty.$$

4. If $a_n = (-1)^n n, n \in \mathbf{N}$, then

$$\underline{\lim}\, a_n = \sup_n \inf \left\{(-1)^n n, (-1)^{n+1}(n+1), \ldots\right\}$$

$$= \sup \{-\infty, -\infty, \ldots\} = -\infty$$

and $$\overline{\lim}\, a_n = +\infty.$$

5. If $a_n = \dfrac{(-1)^n}{n^2}, n \in \mathbf{N}$, then

$$\underline{\lim}\, a_n = \sup_n \inf \left\{\frac{(-1)^n}{n^2}, \frac{(-1)^{n+1}}{(n+1)^2}, \ldots\right\}$$

$$= \sup_n \begin{cases} \dfrac{-1}{n^2}, & \text{if } n \text{ is odd} \\ \dfrac{-1}{(n+1)^2}, & \text{if } n \text{ is even} \end{cases}$$

$$= 0$$

and $$\overline{\lim}\, a_n = 0.$$

6. If $a_n = (-1)^n \left(1+\frac{1}{n}\right), n \in \mathbf{N}$, then

$$\underline{A}_n = \inf\left\{(-1)^n\left(1+\frac{1}{n}\right), (-1)^{n+1}\left(1+\frac{1}{n+1}\right), \ldots\right\}$$

$$= \begin{cases} -\left(1+\frac{1}{n}\right), \text{ if } n \text{ is odd} \\ -\left(1+\frac{1}{n+1}\right), \text{ if } n \text{ is even} \end{cases}$$

$$\underline{\lim}\, a_n = \sup \underline{A}_n = \sup\left\{-2, \frac{-4}{3}, \frac{-6}{5}, \ldots\right\} = -1$$

and $\quad \overline{\lim}\, a_n = 1.$

7. The sequence

$$\left\{-2, 2, \frac{-3}{2}, \frac{3}{2}, \frac{-4}{3}, \frac{4}{3}, \ldots\right\}$$

has limit inferior -1 and limit superior 1. Note that -1 is not lower bound nor 1 is an upper bound of the sequence.

8. If $a_n = n\,(1 + (-1)^n)$, then $\underline{\lim}\, a_n = 0$ and $\overline{\lim}\, a_n = \infty$.

9. If $a_n = \sin\frac{n\pi}{3}, n \in \mathbf{N}$, then the sequence $\{a_n\}$ is

$$\left\{\frac{\sqrt{3}}{2}, \frac{\sqrt{3}}{2}, 0, \frac{-\sqrt{3}}{2}, \frac{-\sqrt{3}}{2}, 0, \frac{\sqrt{3}}{2}, \ldots\right\}$$

$$\underline{A}_n = \frac{-\sqrt{3}}{2} \text{ and } \overline{A}_n = \frac{\sqrt{3}}{2}, \text{ for each } n \in \mathbf{N}$$

$$\underline{\lim}\, a_n = \frac{-\sqrt{3}}{2} \text{ and } \overline{\lim}\, a_n = \frac{\sqrt{3}}{2}$$

10. If $\{r_n\}$ be an enumeration of all the rational numbers between 0 and 1, then

$\underline{\lim}\, r_n = 0$ and $\overline{\lim}\, r_n = 1$

Example 1. If $a_n = \sin\frac{n\pi}{2} + \frac{(-1)^n}{n}, n \in \mathbf{N}$, then show that

$$\underline{\lim}\, a_n = -1 \text{ and } \overline{\lim}\, a_n = 1$$

■ The sequence $\{a_n\}$ is bounded, for $\frac{-4}{3} \le a_n \le 1, \forall\, n$, the terms of the sequence are given by

$$a_{2n} = \frac{1}{2n}, \text{ for all } n,$$

$$a_{4n+1} = 1 - \frac{1}{4n+1}, \text{ for all } n,$$

$$a_{4n+3} = -1 - \frac{1}{4n+3}, \text{ for all } n.$$

Therefore, $\overline{A}_n = 1$ for all n, and so

$$\overline{\lim}\, a_n = 1$$

Also $\underline{A}_n$ is given by

$$\underline{A}_{4n} = -1 - \frac{1}{4n+3}$$

$$\underline{A}_{4n+1} = -1 - \frac{1}{4n+3}$$

$$\underline{A}_{4n+2} = -1 - \frac{1}{4n+3}$$

and $$\underline{A}_{4n+3} = -1 - \frac{1}{4n+3}, \text{for each } n \in \mathrm{N}.$$

Hence $\underline{\lim}\, a_n = -1$

Theorem 7. *If $\{a_n\}$ is any sequence, then*

$\underline{\lim}\, a_n = -\infty$ *if and only if $\{a_n\}$ is not bounded below,*

and $\overline{\lim}\, a_n = +\infty$ *if and only if $\{a_n\}$ is not bounded above.*

Let $\underline{A}_n = \inf \{a_n, a_{n+1}, \ldots\}$,

and $\overline{A}_n = \sup \{a_n, a_{n+1}, \ldots\}, n \in \mathbf{N}$

By definition, we have

$$\underline{\lim}\, a_n = -\infty \Leftrightarrow \sup \{\underline{A}_1, \underline{A}_2, \ldots\} = -\infty$$

$\Leftrightarrow$ $\underline{A}_n = -\infty, \forall\, n \in \mathbf{N} \Leftrightarrow \underline{A}_n = -\infty, \forall\, n \in \mathbf{N}$

$\Leftrightarrow$ $\inf \{a_n, a_{n+1}, \ldots\} = -\infty, \forall\, n \in \mathbf{N}$

$\Leftrightarrow$ $\{a_n\}$ is not bounded below.

The proof for limit superior is similar.

Corollary. If $\{a_n\}$ is any sequence, then

(*i*) $-\infty < \underline{\lim}\, a_n \le +\infty$ iff $\{a_n\}$ is bounded below.

and (*ii*) $-\infty \le \overline{\lim}\, a_n < +\infty$ iff $\{a_n\}$ is bounded above.

For bounded sequences, we have the following useful criteria for limits inferior and superior respectively.

Theorem 8. *A real number $\underline{a}$ is the limit inferior of a bounded sequence $\{a_n\}$ iff for each $\varepsilon > 0$, the following results hold:*

(*i*) $a_n < \underline{a} + \varepsilon$, *for infinitely many values of n, and*

(*ii*) *there exists a positive integer m such that*

$$a_n > \underline{a} - \varepsilon, \text{ for all } n \geq m.$$

Since $\{a_n\}$ is bounded, therefore $\underline{\lim}\, a_n$ is finite. Let $\varepsilon > 0$ be given. Then, we have

$$\underline{a} = \lim a_n = \sup_n \underline{A}_n \Leftrightarrow \begin{cases} \underline{A}_n \leq \underline{a}, \text{ for all } n, \\ \exists\, m \in \mathbf{N} \text{ such that} \\ \underline{A}_m > \underline{a} - \varepsilon \end{cases}$$

$$\Leftrightarrow \begin{cases} \inf\{a_n, a_{n+1}, \ldots\} \leq \underline{a}, \forall n, \\ \exists\, m \in \mathbf{N} \text{ such that} \\ \inf\{a_m, a_{m+1}, \ldots\} > \underline{a} - \varepsilon \end{cases}$$

$$\Leftrightarrow \begin{cases} a_n < \underline{a} + \varepsilon, \text{ for infinitely many values of } n, \\ \exists\, m \in \mathbf{N} \text{ such that} \\ a_n > \underline{a} - \varepsilon, \forall\, n \geq m. \end{cases}$$

Hence the result follows.

Corollary 1. If $\{a_n\}$ is a bounded sequence and $\underline{a} = (\underline{\lim}\, a_n,)$, then

(*i*) for each real number $\alpha < \underline{a}$, $\exists$ a positive integer m such that

$$a_n > \alpha, \quad \forall\, n \geq m, \text{ and}$$

(*ii*) if $\alpha \in \mathbf{R}$ and $\exists\, m \in \mathbf{N}$ such that $a_n > \alpha$, $\forall\, n \geq m$, then $\underline{\lim}\, a_n \geq \alpha$.

For (*i*), take $\varepsilon = a - \alpha$ and use first part of the above theorem, and

for (*ii*), such an α is then a lower bound of the set $\{a_n : n \geq m\}$

and so $$\underline{\lim}\, a_n = \sup_n \underline{A}_n \geq \underline{A}_m = \inf\{a_n : n \geq m\} \geq \alpha.$$

Corollary 2. Let $\{a_n\}$ be a bounded sequence, then

(*i*) if $\alpha \in \mathbf{R}$ and $\underline{\lim}\, a_n < \alpha$, then $a_n < \alpha$, for infinitely many values of n, and

(*ii*) if $\alpha \in \mathbf{R}$ is such that $\{n : a_n < \alpha\}$ is infinite, then $\underline{\lim}\, a_n \leq \alpha$.

The following theorem follows by applying the above theorem to the sequence $\{-a_n\}$ and using theorem (6 of section 3).

Theorem 9. *A real number is $\overline{a}$ the limit superior of a bounded sequence $\{a_n\}$ iff for each $\varepsilon > 0$, the following results hold:*

(*i*) $a_n > \overline{a} - \varepsilon$, *for infinitely many values of n,*

and (*ii*) $\exists$ *a positive integer m such that*

$$a_n < \overline{a} + \varepsilon, \quad \forall\, n \geq m.$$

CHAPTER 3

Corollary 1. If $\{a_n\}$ is a bounded sequence and $\bar{a} = \overline{\lim}\, a_n$, then

(*i*) for each real number $\beta > \bar{a}$, $\exists\, m \in \mathbf{N}$, such that $a_n < \beta$, $\forall\, n \geq m$, and

(*ii*) if $\beta \in \mathbf{R}$ and $\exists\, m \in \mathbf{N}$ such that $a_n < \beta$, $\forall\, n \geq m$, then $\overline{\lim}\, a_n \leq \beta$.

Corollary 2. Let $\{a_n\}$ be a bounded sequence and $\beta \in \mathbf{R}$.

(*i*) if $\overline{\lim}\, a_n > \beta$, then $\{n : a_n > \beta\}$ is infinite, and

(*ii*) if $\{n : a_n > \beta\}$ is infinite, then $\overline{\lim}\, a_n \geq \beta$.

The following theorem shows that the limits–inferior and superior are the *smallest* and the *greatest* limit points, respectively, of a bounded sequence and hence the *lower* and *upper* limits of the bounded sequence.

Theorem 10. If $\{a_n\}$ *is a bounded sequence, then*

(*i*) $\underline{\lim}\, a_n$ *= smallest limit point of* $\{a_n\}$, *and*

(*ii*) $\overline{\lim}\, a_n$ *= greatest limit point of* $\{a_n\}$.

We shall prove (*i*) and leave (*ii*) as an exercise to the reader.

Let $\underline{\lim}\, a_n = \underline{a}$. Then $\underline{a}$ is finite, since $\{a_n\}$ is bounded.

Now by theorem (8), it follows that for each $\varepsilon > 0$,

$$\underline{a} - \varepsilon < a_n < \underline{a} + \varepsilon, \text{ for infinitely many values of } n.$$

Thus $\underline{a}$ is a limit point of $\{a_n\}$. Moreover, by the second part of the theorem (8), we have, for any $\varepsilon > 0$, there are at the most a finite number of terms of the bounded sequence $\{a_n\}$, for which $a_n \leq \underline{a} - \varepsilon$, and consequently any number smaller than $\underline{a}$ is not a limit point of $\{a_n\}$. Hence, $a = \underline{\lim}\, a_n$ is the smallest limit point of the bounded sequence $\{a_n\}$.

4. CONVERGENT SEQUENCES

A sequence may have no limit point, a unique limit point or a finite or an infinite number of limit points. Our interest lies chiefly in a bounded sequence with a unique limit point. Evidently such a sequence can have at the most a finite number of terms outside the interval $]l - \varepsilon, l + \varepsilon[$, $\varepsilon > 0$, however, small ε may be. For otherwise, by the *Bolzano-Weierstrass Theorem* the infinite number of outside terms will have another limit point. Further, the condition automatically ensures the existence of an infinite number of terms of the sequence within the interval. Such sequences are called *convergent sequences*.

Let us proceed to show that the converse of theorem 3 also holds.

4.1 Theorem 11. *Every bounded sequence with a unique limit point is convergent.*

Let l be the only limit point of a bounded sequence $\{S_n\}$, which surely exists by *Bolzano-Weierstrass Theorem*. Thus, for $\varepsilon > 0$, $S_n \in\,]l - \varepsilon, l + \varepsilon[$ for an infinite number of values of $n \in \mathbf{N}$.

l being the only limit point of the sequence there can exist only a finite number of values say m_1, m_2, ..., m_r of n such that the corresponding terms of the sequence do not belong to $]l - \varepsilon, l + \varepsilon[$. For otherwise, the infinitely many outside terms will have a limit point other than l.

Let $(m-1)$ be the greatest of such exceptional values of n. Thus, we have

$$S_n \in]l-\varepsilon, l+\varepsilon[, \quad \forall\, n \geq m,$$

i.e.,

$$|S_n - l| < \varepsilon, \quad \forall\, n \geq m.$$

Thus the sequence $\{S_n\}$ converges to its unique limit point l.

Theorems 3 and 11 may be stated in the combined form as:

Theorem 12. *A necessary and sufficient condition for the convergence of a sequence is that it is bounded and has a unique limit point.*

4.2 In view of the above discussion, we give below another equivalent definition of a convergent sequence.

Definition 2. A sequence is said to be convergent if it is bounded and has a unique limit point.

Starting with the definition 2 for the convergence of a sequence the reader is advised to prove the following theorem.

Theorem 13. *A necessary and sufficient condition for a sequence $\{S_n\}$ to converge to l is that to each $\varepsilon > 0$, there corresponds a positive integer m such that*

$$|S_n - l| < \varepsilon, \quad \forall\, n \geq m.$$

The following theorem shows that a bounded sequence converges iff its limits inferior and superior coincide and the common value is the limit of the sequence.

Theorem 14. *A bounded sequence $\{a_n\}$ converges to a real number a if and only if*

$$\underline{\lim}\, a_n = \overline{\lim}\, a_n = a.$$

The condition is necessary: If the sequence $\{a_n\}$ converges to a, then, by theorem (12), a is the unique limit point of $\{a_n\}$. Hence the limits inferior and limit superior both are equal to a (using Theorem (10)).

The condition is sufficient: Let $\{a_n\}$ be a bounded sequence such that

$$\underline{\lim}\, a_n = \overline{\lim}\, a_n = a.$$

This shows that a is the unique limit point of the bounded sequence $\{a_n\}$. Since limit inferior and limit superior are the smallest and greatest limit points. Hence by theorem 12, it follows that $\{a_n\}$ converges to a.

5. NON-CONVERGENT SEQUENCES (DEFINITIONS)

(*a*) *Bounded Sequences*

A bounded sequence which does not converge, and has at least two limit points, is said to **oscillate finitely.**

(*b*) *Unbounded Sequences*

(*i*) If a sequence $\{S_n\}$ is unbounded on the left (below), then we say that $-\infty$ is a limit point of the sequence, and to each positive number G, however large, there corresponds a positive integer m, such that

$$S_n < -G, \quad \forall\, n \geq m,$$

i.e., the sequence has an infinity of terms below $-G$.

Also, then $-\infty$ is the least limit point so that

$$\underline{\lim}\, S_n = -\infty.$$

(*ii*) If a sequence $\{S_n\}$ is unbounded on the right (above), then we say that $+\infty$ is a limit point of the sequence, and to each positive number G, however large, there corresponds a positive integer m, such that

$$S_n > G, \quad \forall\, n \geq m,$$

i.e., the sequence has an infinite number of terms above G. Also, then $+\infty$ is the greatest limit point, so that $\overline{\lim}\, S_n = +\infty$.

(*iii*) If a sequence $\{S_n\}$ is bounded above (on the right) but not below and besides $-\infty$, has *no other* limit point, then $-\infty$ is not only its least but also its greatest limit point, and so we equate the upper limit also to $-\infty$, *i.e.*,

$$\underline{\lim}\, S_n = \overline{\lim}\, S_n = \lim S_n = -\infty.$$

The sequence is then said to **diverge to** $-\infty$.

(*iv*) If, finally, the sequence is bounded on the left (below) but not on the right (above) and besides $+\infty$, has *no other* limit point, then $+\infty$ is not only its greatest but also its least limit point. So we have

$$\overline{\lim}\, S_n = \underline{\lim}\, S_n = \lim S_n = +\infty.$$

The sequence is, then, said to **diverge to** $+\infty$.

(*v*) An unbounded sequence is said to oscillate infinitely if it diverges neither to $+\infty$ nor to $-\infty$. Thus a bounded sequence either *converges* or else *oscillates finitely*, but an unbounded sequence either *diverges* to $+\infty$ or $-\infty$ or *oscillates infinitely*.

Note: ∞ is not considered here as a real number because we are not dealing with the extended real number system. In the latter case, the definitions need lot of modifications.

ILLUSTRATIONS

1. $\{1 + (-1)^n\}$ oscillates finitely.
2. $\left\{(-1)^n\left(1 + \dfrac{1}{n}\right)\right\}$ oscillates finitely.
3. $\{n^2\}$ diverges to $+\infty$.
4. $\{-2^n\}$ diverges to $-\infty$.
5. $\{n(-1)^n\}$ oscillates infinitely.
6. $\left\{\dfrac{(-1)^{n-1}}{n!}\right\}$ converges to the limit 0.

7. $\left\{1+\frac{1}{n}\right\}$ converges to the limit 1.

8. $\left\{1, 2, \frac{1}{2}, 3, \frac{1}{3}, \ldots\right\}$ is bounded below but unbounded above, and has a limit point 0 besides $+\infty$,

$$\underline{\lim}\, S_n = 0,\ \overline{\lim}\, S_n = +\infty.$$

The sequence oscillates infinitely.

9. $\{1, 2, 3, 2, 5, 2, 7, 2, 3, 2, 11, 2, 13, \ldots\}$,

where, $S_n = \begin{cases} 2, \text{ when } n \text{ is even,} \\ \text{lowest prime factor } (\neq 1) \text{ of } n, \text{ when } n \text{ is odd,} \end{cases}$

is bounded on the left but not on the right. It has infinite number of limit points 2, 3, 5, 7, 11, ..., so that

$$\underline{\lim}\, S_n = 2,\ \overline{\lim}\, S_n = +\infty.$$

The sequence oscillates infinitely.

10. The sequence $\left\{m+\frac{1}{n}\right\}$ where m, n are natural numbers, also oscillates infinitely 1, 2, 3, ... being its limit points.

Solved Examples

Example 2. Show that $\lim_{n\to\infty} \frac{3+2\sqrt{n}}{\sqrt{n}} = 2$.

■ Let ε be any positive number.

$$\therefore \quad \left|\frac{3+2\sqrt{n}}{\sqrt{n}} - 2\right| < \varepsilon, \text{ if } \left|\frac{3}{\sqrt{n}}\right| < \varepsilon \text{ or } n > \frac{9}{\varepsilon^2}.$$

Let m be a positive integer greater than $9/\varepsilon^2$.

Thus to $\varepsilon > 0, \exists$ a positive m, such that

$$\left|\frac{3+2\sqrt{n}}{\sqrt{n}} - 2\right| < \varepsilon,\ \forall\, n \geq m$$

$$\therefore \quad \lim_{n\to\infty} \frac{3+2\sqrt{n}}{\sqrt{n}} = 2.$$

Example 3. Show that

(*i*) $\lim_{n\to\infty} \sqrt[n]{a} = 1$, if $a > 0$ and

(*ii*) $\lim_{n\to\infty} \sqrt[n]{n} = 1$

■ (*i*) If $a > 1$, let $\sqrt[n]{a} = 1 + h_n$, where $h_n > 0$

$$\therefore \quad a = (1 + h_n)^n = 1 + n\,h_n + \frac{n\,(n-1)}{1.2} h_{n^2} + \ldots + h_{n^n}$$

$$> 1 + n\,h_n, \ \forall\, n \qquad (\because h_n > 0)$$

$$\therefore \quad h_n < \frac{a-1}{n}, \ \forall\, n$$

Let ε be any positive number, then

$$|h_n| = h_n < \frac{a-1}{n} < \varepsilon, \text{ where } n > \frac{a-1}{\varepsilon}$$

Let m be any positive integer $> \dfrac{a-1}{\varepsilon}$, then

$$\left|\sqrt[n]{a} - 1\right| = |h_n| < \varepsilon, \ \forall\, n \geq m.$$

If $a = 1$, the result is trivial and if $0 < a < 1$ the result is obtained by taking reciprocals.

(*ii*) Let $\sqrt[n]{n} = 1 + h_n$, where $h_n \geq 0$

$$\therefore \quad n = (1 + h_n)^n = 1 + nh_n + \frac{1}{2} n(n-1)h_n^2 + \ldots + h_n^n$$

$$> \frac{1}{2} n(n-1)h_n^2, \ \forall n \qquad (\because h_n \geq 0)$$

$$\therefore \quad h_n^2 < \frac{2}{n-1}, \text{ for } n \geq 2$$

$$\text{or} \quad |h_n| < \sqrt{\frac{2}{n-1}}, \text{ for } n \geq 2.$$

Let ε be any positive number, then

$$|h_n| < \sqrt{\frac{2}{n-1}} < \varepsilon, \text{ when } n > 1 + 2/\varepsilon^2.$$

Let m be any positive integer $> 1 + 2/\varepsilon^2$.

Thus for $\varepsilon > 0, \exists$ a positive integer m such that

$$\left|\sqrt[n]{n} - 1\right| = |h_n| < \varepsilon, \ \forall\, n \geq m.$$

Hence, $\quad \lim_{n\to\infty} \sqrt[n]{n} = 1.$

Example 4. Show that the sequence $\{r^n\}$ converges iff $-1 < r \leq 1$.

■ (*i*) When $r > 1$, let $r = 1 + h$, $h > 0$.

$$\therefore \qquad r^n = (1+h)^n$$
$$> 1 + nh, \quad \forall\, n \in \mathbf{N}$$

If $G > 0$ be any number however large, we have

$$1 + nh > G, \text{ if } n > \frac{G-1}{h}.$$

Let m be a positive integer greater than $\dfrac{G-1}{h}$,

$\therefore$ for $G > 0$, $\exists$ a positive integer m such that

$$r^n > G, \quad \forall\, n \geq m.$$

Hence, the sequence diverges to ∞.

(*ii*) When $r = 1$, evidently $\lim r^n = 1$.

$\therefore$ The sequence converges to 1.

(*iii*) When $|r| < 1$, let $|r| = \dfrac{1}{1+h}$, where $h > 0$.

$$\therefore \qquad |r^n| = |r|^n = \frac{1}{(1+h)^n} \leq \frac{1}{1+nh}, \quad \forall\, n \in \mathbf{N}.$$

Let ε be any positive number, then

$$\frac{1}{1+nh} < \varepsilon, \text{ when } n > \left(\frac{1}{\varepsilon} - 1\right)\Big/ h$$

Let m be a positive integer greater than $\left(\dfrac{1}{\varepsilon} - 1\right)\Big/ h$,

$\therefore$ for $\varepsilon > 0$, $\exists$ a positive integer m such that

$$|r^n| < \varepsilon, \quad \forall\, n \geq m.$$

Hence, $\{r^n\}$ converges to 0, *i.e.*,

$$\lim r^n = 0, \text{ when } |r| < 1.$$

(*iv*) When $r = -1$, the sequence $\{(-1)^n\}$ is bounded and has two limit points.

$\therefore$ the sequence oscillates finitely.

(*v*) When $r < -1$, let $r = -t$ so that $t > 1$.

Thus we get the sequence $\{(-1)^n t^n\}$, which has both positive and negative terms. The sequence is unbounded and the numerical values of the terms can be made greater than any number however large. Thus, it oscillates infinitely.

Hence, the sequence $\{r^n\}$ converges only when $-1 < r \leq 1$.

Note: The sequence $\{r^n\}$ converges to zero iff $|r| < 1$.

Example 5. If $a > 0$ and p is real, then $\lim\limits_{n\to\infty} \dfrac{n^p}{(1+a)^n} = 0$.

■ Let $r > p$ be any positive integer. For $n > 2r$

$$(1+a)^n > {}^nc_r\, a^r = \frac{n(n-1)\dots(n-r+1)}{\lambda_r} a^r > \left(\frac{n}{2}\right)^r \frac{a^r}{\lambda_r}$$

Hence $$0 < \frac{n^p}{(1+a)^n} < \frac{2^r \lambda_r n^p}{a^r n^r} = \frac{2^r \lambda_r n^{p-r}}{a^r}$$

Since $p - r < 0$, therefore given $\varepsilon > 0$, $\exists$ a positive integer m such that

$$\left|n^{p-r} - 0\right| < \varepsilon, \text{ whenever } n \geq m \left[m > \frac{1}{\varepsilon^{1/r-p}}\right]$$

Thus $\lim\limits_{n\to\infty} \dfrac{n^p}{(1+a)^n} = 0$.

Example 4 follows from this if we take $p = 0$.

EXERCISE

Show that:

1. $\lim \dfrac{2n-3}{n+1} = 2$
2. The sequence $\{(-1)^n\}$ oscillates finitely.
3. The sequence $\{S_n\}$, where $S_n = 1 + \dfrac{(-1)^n}{n}$ converges.
4. $\lim \dfrac{1+2+3+\dots+n}{n^2} = \dfrac{1}{2}$
5. $\lim \dfrac{1+3+5+\dots+(2n-1)}{n^2} = 1$.
6. The sequence $\{n + (-1)^n n\}$ oscillates infinitely.

6. CAUCHY'S GENERAL PRINCIPLE OF CONVERGENCE

Theorem 13 can be used to test the convergence of a sequence to a given limit l, but in cases where limit l is not known, nor can any guess be made of the same, the following theorem which involves only the terms of the sequence, is useful for determining whether a sequence converges or not.

Theorem 15. *A necessary and sufficient condition for the convergence of a sequence $\{S_n\}$ is that, for each $\varepsilon > 0$ there exists a positive integer m such that*

$$\left|S_{n+p} - S_n\right| < \varepsilon, \quad \forall\, n \geq m \wedge p \geq 1.$$

Necessary: Let the sequence be convergent and let l be its limit, so that for a given $\varepsilon > 0$, $\exists$ a positive integer m, such that

$$|S_n - l| < \frac{1}{2}\varepsilon, \ \forall n \geq m$$

If $p \geq 1$, then $n + p > n \geq m$, and so

$$|S_{n+p} - l| < \frac{1}{2}\varepsilon, \ \forall n \geq m \wedge p \geq 1$$

$$\Rightarrow \quad |S_{n+p} - S_n| = |S_{n+p} - l + l - S_n|$$

$$\leq |S_{n+p} - l| + |l - S_n|$$

$$< \frac{1}{2}\varepsilon + \frac{1}{2}\varepsilon = \varepsilon, \ \forall n \geq m \wedge p \geq 1.$$

Sufficient: To establish the convergence of the sequence as a consequence of the given conditions, we first show that the sequence is bounded and then, that it converges to a limit.

Now by the given condition for $\varepsilon = 1$, $\exists$ a positive integer m_0 such that

$$|S_{n+p} - S_n| < 1, \ \forall \, n \geq m_0 \wedge p \geq 1$$

In particular, for $n = m_0$

$$|S_{m_0+p} - S_{m_0}| < 1, \ \forall \, p \geq 1$$

i.e.,

$$S_{m_0} - 1 < S_{m_0+p} < S_{m_0} + 1, \ \forall \, p \geq 1.$$

Let $g = \min\{S_1, S_2, ..., S_{m_0-1}, S_{m_0} - 1\}$

$G = \max\{S_1, S_2, ..., S_{m_0-1}, S_{m_0} + 1\}$.

Then, $g \leq S_n \leq G, \ \forall \, n$

Hence, the sequence is bounded and therefore by *Bolzano-Weierstrass Theorem* for sequences, it has at least one limit point, say l. We shall now show that the sequence converges to l, *i.e.*, $\lim S_n = l$.

Now by the given condition, for $\varepsilon > 0$, $\exists$ a positive integer m such that

$$|S_{n+p} - S_n| < \frac{1}{3}\varepsilon, \ \forall n \geq m \wedge p \geq 1$$

In particular, for $n = m$

$$|S_{m+p} - S_m| < \frac{1}{3}\varepsilon, \ \forall p \geq 1 \qquad ...(1)$$

As l is a limit point, $\exists$ an integer $m_1 > m$ such that

$$|S_{m_1} - l| < \frac{1}{3}\varepsilon \qquad ...(2)$$

Also, since $m_1 > m$ therefore from (1), we have

$$|S_{m_1} - S_m| < \frac{1}{3}\varepsilon \qquad ...(3)$$

CHAPTER 3

$$\therefore \quad \left|S_{m+p} - l\right| = \left|S_{m+p} - S_m + S_m - S_{m_1} + S_{m_1} - l\right|$$

$$\leq \left|S_{m+p} - S_m\right| + \left|S_m - S_{m_1}\right| + \left|S_{m_1} - l\right|$$

$$< \frac{1}{3}\varepsilon + \frac{1}{3}\varepsilon + \frac{1}{3}\varepsilon = \varepsilon, \forall\, p \geq 1$$

$$\Rightarrow \quad \left|S_n - l\right| < \varepsilon, \ \forall\, n \geq m$$

Hence, the sequence $\{S_n\}$ converges to l.

6.1 Cauchy Sequence

A sequence $\{S_n\}$ is called a *Cauchy sequence* or a *fundamental sequence* if for each $\varepsilon > 0$, $\exists$ a positive integer m, such that

$$\left|S_{n+p} - S_n\right| < \varepsilon, \ \forall\, n \geq m \wedge p \geq 1$$

or

$$\left|S_{n_1} - S_{n_2}\right| < \varepsilon, \ \forall\, n_1, n_2 \geq m$$

Thus in the field of real numbers, a sequence is convergent iff it is a Cauchy sequence.

Note: A sequence cannot converge if even one $\varepsilon > 0$ can be found such that for every positive integer m,

$$\left|S_{n+p} - S_n\right| \nless \varepsilon, \ \forall\, n \geq m \wedge p \geq 1$$

Ex. 1. If $\{a_n\}$ and $\{b_n\}$ are two Cauchy sequences, then the sequences $\{a_n \pm b_n\}$, $\{a_n \cdot b_n\}$ and $\{a_n/b_n\}$ (if $b_n \neq 0$ for all n) are also Cauchy sequences.

Ex. 2. Show that every Cauchy sequence is bounded. Is the converse true?

Example 6. Show that the sequence $\{S_n\}$, where $S_n = 1 + \frac{1}{2} + \frac{1}{3} + \ldots + \frac{1}{n}$ cannot converge.

■ Suppose, if possible, the sequence $\{S_n\}$ is convergent.

Let us take $\varepsilon = \frac{1}{2}$, and $n = m$ and $p = m$ in Cauchy's General Principle of Convergence, so that

$$\left|S_{2m} - S_m\right| < \frac{1}{2}$$

$$\text{But } S_{2m} - S_m = \frac{1}{m-1} + \frac{1}{m-2} + \ldots + \frac{1}{2m},$$

$$> \frac{1}{2m} + \frac{1}{2m} + \ldots + \frac{1}{2m} = \frac{m}{2m} = \frac{1}{2}$$

i.e., $\left|S_{2m} - S_m\right| > \frac{1}{2}$, which contradicts Cauchy's general principle of convergence.

Hence, the sequence cannot converge.

Ex. Show that the sequence $\{S_n\}$, where

(*i*) $S_n = 1 + \dfrac{1}{3} + \dfrac{1}{5} + \ldots + \dfrac{1}{2n - 1}$

(*ii*) $S_n = 1 + \dfrac{1}{4} + \dfrac{1}{7} + \ldots + \dfrac{1}{3n - 2}$

cannot converge.

7. ALGEBRA OF SEQUENCES

It may not always be easy to prove the convergence of a sequence by direct application of the definition of convergence or by Cauchy's General Principle of Convergence. But it may enable us to evaluate the limit of a sequence whose terms can be expressed as a sum, difference, product or quotient of the corresponding terms of two convergent sequences. The sequence $\{S_n\}$ whose nth term is $a_n + b_n$, $a_n - b_n$, $a_n b_n$ or $\dfrac{a_n}{b_n}$ $(b_n \neq 0)$ is called the sum, difference, product or quotient of the sequences $\{a_n\}$ and $\{b_n\}$. We shall now discuss the convergence of such sequences.

Theorem 16. *If* $\{a_n\}, \{b_n\}$ *be two sequences such that* $\lim a_n = a$, $\lim b_n = b$, *then*

(*i*) $\lim (a_n \pm b_n) = \lim a_n \pm \lim b_n = a \pm b$

(*ii*) $\lim (a_n b_n) = (\lim a_n)(\lim b_n) = ab$

(*iii*) $\lim (a_n / b_n) = (\lim a_n / \lim b_n)$, if $b \neq 0, b_n \neq 0, \ \forall\, n$.

(*i*) Let $\varepsilon > 0$ be given.

Since $\lim a_n = a$ and $\lim b_n = b$,

$\therefore$ $\exists$ positive integers m_1 and m_2, respectively, such that

$$|a_n - a| < \frac{1}{2}\varepsilon, \ \forall\, n \geq m_1, \text{ and}$$

$$|b_n - b| < \frac{1}{2}\varepsilon, \ \forall\, n \geq m_2.$$

Thus, for $m = \max(m_1, m_2)$, we have

$$|a_n - a| < \frac{1}{2}\varepsilon, \ |b_n - b| < \frac{1}{2}\varepsilon, \ \forall\, n \geq m$$

$$\therefore \quad |(a_n \pm b_n) - (a \pm b)| = |(a_n - a) \pm (b_n - b)|$$
$$\leq |a_n - a| \pm |b_n - b|$$
$$< \frac{1}{2}\varepsilon + \frac{1}{2}\varepsilon - \varepsilon, \ \forall\, n \geq m$$

Hence, $\quad \lim (a_n \pm b_n) = a \pm b - \lim a_n \pm \lim b_n$.

(*ii*) The two sequences $\{a_n\}, \{b_n\}$ being convergent, are bounded so that $\exists$ positive real numbers k, K such that

$$|a_n| \leq k, |b_n| \leq K, \ \forall\, n$$

Now

$$\left|a_nb_n - ab\right| = \left|a_n(b_n - b) + b(a_n - a)\right|$$

$$\leq \left|a_n\right| \cdot \left|b_n - b\right| + \left|b\right| \cdot \left|a_n - a\right|$$

$$\leq k\left|b_n - b\right| + \left|b\right| \cdot \left|a_n - a\right| \qquad ...(1)$$

Let $\varepsilon > 0$ be given.

Since $\lim a_n = a$, $\lim b_n = b$, therefore there exist positive integers m_1 and m_2, respectively, such that

$$\left|a_n - a\right| < \frac{\frac{1}{2}\varepsilon}{\left|b\right| + 1}, \ \forall\, n \geq m_1$$

$$\left|b_n - b\right| < \frac{\frac{1}{2}\varepsilon}{k}, \ \forall\, n \geq m_2$$

Then, for $m = \max\,(m_1, m_2)$, we have from (1)

$$\left|a_nb_n - ab\right| < \frac{1}{2}\varepsilon + \frac{\left|b\right| \cdot \frac{1}{2}\varepsilon}{\left|b\right| + 1} < \varepsilon, \ \forall\, n \geq m.$$

Hence, $\qquad \lim\,(a_nb_n) = ab = (\lim a_n)\,(\lim b_n).$

(*iii*) *Lemma*. To show that if $\lim b_n = b \neq 0$, then $\exists$ a positive number λ and a positive integer m_3 such that

$$\left|b_n\right| > \lambda, \ \forall\, n \geq m_3$$

Let us take $\varepsilon = \frac{1}{2}\left|b\right|$, so that there exists a positive integer m_3 such that

$$\left|b_n - b\right| < \frac{1}{2}\left|b\right|, \ \forall\, n \geq m_3,$$

Thus, $$\left|b\right| - \left|b_n\right| \geq \left|b_n - b\right| < \frac{1}{2}\left|b\right|.$$

$\Rightarrow$ $$\left|b_n\right| \geq \frac{1}{2}\left|b\right| \text{ (say)}, \ \forall\, n \geq m_3. \qquad ...(2)$$

Let us apply the Lemma to prove the main theorem.

Now

$$\left|\frac{a_n}{b_n} - \frac{a}{b}\right| = \left|\frac{ba_n - ab_n}{bb_n}\right| = \left|\frac{b(a_n - a) - a\,(b_n - b)}{bb_n}\right|$$

$$\leq \frac{\left|b\right|\left|a_n - a\right| + \left|a\right|\left|b_n - b\right|}{\left|b\right|\left|b_n\right|}$$

$$\leq \frac{2}{\left|b\right|}\left|a_n - a\right| + \frac{2\left|a\right|}{\left|b\right|^2}\left|b_n - b\right|, \ \forall\, n \geq m_3 \qquad \text{[using (2)]}$$

Let $\varepsilon > 0$ be given.

Since $\lim a_n = a$, $\lim b_n = b$, therefore, $\exists$ positive integers m_1, m_2 such that

$$|a_n - a| < \frac{1}{4}|b|\varepsilon, \ \forall\, n \geq m_1$$

and
$$|b_n - b| < \frac{1}{4}\frac{|b|^2 \varepsilon}{|a|+1}, \ \forall\, n \geq m_2.$$

Thus, for $m = \max(m_1, m_2, m_3)$, we have

$$\left|\frac{a_n}{b_n} - \frac{a}{b}\right| < \frac{1}{2}\varepsilon + \frac{1}{2}\varepsilon = \varepsilon, \ \forall\, n \geq m$$

Hence,
$$\lim\left(\frac{a_n}{b_n}\right) = \frac{a}{b} = \frac{\lim a_n}{\lim b_n}.$$

Note: It may be noted that while the sum, difference, product and quotient under certain conditions, of two convergent sequences is convergent, the converse may not be true, *i.e.*, if the sequence $\{a_n \pm b_n\}$, $\{a_n b_n\}$ or $\left\{\frac{a_n}{b_n}\right\}$ is convergent, the component sequences $\{a_n\}$ and $\{b_n\}$ may not be convergent, however, it is not difficult to see that both shall behave alike.

For example, consider the sequences $\{a_n\}$ and $\{b_n\}$.

(1) When $a_n = n^2$ and $b_n = -n^2$.

The sequence $\{a_n + b_n\}$ converges to zero and the sequence $\left\{\frac{a_n}{b_n}\right\}$ converges to -1, but the two sequences $\{a_n\}$, $\{b_n\}$ are divergent.

(2) If $a_n = b_n = (-1)^n$.

The sequence $\{a_n - b_n\}$ converges to zero, $\left\{\frac{a_n}{b_n}\right\}$ converges to 1 and $\{a_n b_n\}$ converges to 1, while both the sequences $\{a_n\}$ and $\{b_n\}$ oscillate finitely.

(3) If $a_n = (-1)^n$ and $b_n = (-1)^{n+1}$.

The sequence $\{a_n + b_n\}$ converges to zero, $\{a_n b_n\}$ converges to -1 and $\left\{\frac{a_n}{b_n}\right\}$ converges to -1, but $\{a_n\}$ and $\{b_n\}$ are not convergent.

There are many inequalities relating the inferior and superior limits to the arithmetic operations of sequences. We shall now prove some of them, as a sample of the technique involved, for bounded sequences only. The results are, however, true for unbounded sequences as well with well-defined arithmetic operations.

Theorem 17. *If $\{a_n\}$ and $\{b_n\}$ are bounded sequences, then*

(*i*) $\underline{\lim}\, a_n + \underline{\lim}\, b_n \leq \underline{\lim}\,(a_n + b_n)$,

(*ii*) $\underline{\lim}\,(a_n + b_n) \leq \underline{\lim}\, a_n + \overline{\lim}\, b_n$,

CHAPTER 3

(*iii*) $\underline{\lim}\, a_n + \overline{\lim}\, b_n \le \overline{\lim}\,(a_n + b_n)$,

(*iv*) $\overline{\lim}\,(a_n + b_n) \le \overline{\lim}\, a_n + \overline{\lim}\, b_n$.

We shall prove only (*i*) and (*iii*). The other inequalities follow from (*i*) and (*iii*) applying to the sequences $\{-a_n\}$ and $\{-b_n\}$ respectively and using theorem (6 of section 3).

(*i*) Let

$$\underline{\lim}\, a_n = \underline{a} \text{ and } \underline{\lim}\, b_n = \underline{b},$$

then both $\underline{a}$ and $\underline{b}$ are real numbers, since $\{a_n\}$ and $\{b_n\}$ are bounded sequences. Let $\varepsilon > 0$ be given. Then by theorem (8 (*ii*) of section 3) there exist positive integers m_1 and m_2, respectively, such that

$$a_n > \underline{a} - \varepsilon/2, \quad \forall\, n \ge m_1 \qquad \ldots(1)$$

and

$$b_n > \underline{b} - \varepsilon/2, \quad \forall\, n \ge m_2 \qquad \ldots(2)$$

Put $m = \max(m_1, m_2)$ then from (1) and (2), we have

$$a_n + b_n > \underline{a} + \underline{b} - \varepsilon, \quad \forall\, n \ge m$$

This implies that

$$\underline{a} + \underline{b} - \varepsilon \le \inf_{n \ge m} (a_n + b_n) \le \sup_{m} \inf_{n \ge m} (a_n + b_n) = \underline{\lim}\,(a_n + b_n)$$

But $\varepsilon > 0$ was chosen arbitrarily, hence

$$\underline{\lim}\, a_n + \underline{\lim}\, b_n = \underline{a} + \underline{b} \le \underline{\lim}\,(a_n + b_n)$$

(*iii*) Let

$$\underline{\lim}\, a_n = \underline{a} \text{ and } \overline{\lim}\, b_n = \bar{b},$$

then $\underline{a}, \bar{b} \in \mathbf{R}$. Let $\varepsilon > 0$ be given. Then by theorem (3.8 (*ii*) of section 3), there exists a positive integer m such that

$$a_n > \underline{a} - \varepsilon/2, \quad \forall\, n \ge m$$

Also by Theorem (9 (*i*) of Section 3), we have

$$b_n > \bar{b} - \varepsilon/2,$$

for infinitely many values of n. Thus

$$a_n + b_n > \underline{a} + \bar{b} - \varepsilon,$$

for infinitely many values of n.

Therefore, for given $k \in \mathbf{N}, \exists\, n_0 \ge k$ such that

$$a_{n_0} + b_{n_0} > \underline{a} + \bar{b} - \varepsilon$$

This implies that

$$\sup_{n \ge k} (a_n + b_n) \ge a_{n_0} + b_{n_0} > \underline{a} + \bar{b} - \varepsilon,$$

for each $k \in \mathbf{N}$.

Hence, $\underline{a} + \bar{b} - \varepsilon \le \inf_k \sup_{n \ge k} (a_n + b_n) = \overline{\lim}\,(a_n + b_n)$.

The inequality (*iii*) follows, since $\varepsilon > 0$ was arbitrary.

Corollary. If $\{a_n\}$ and $\{b_n\}$ are bounded sequences, then

$$\underline{\lim}\, a_n - \overline{\lim}\, b_n \le \underline{\lim}\,(a_n - b_n) \le \begin{Bmatrix} \overline{\lim}\, a_n - \overline{\lim}\, b_n \\ \underline{\lim}\, a_n - \underline{\lim}\, b_n \end{Bmatrix}$$

$$\le \overline{\lim}\,(a_n - b_n) \le \overline{\lim}\, a_n - \underline{\lim}\, b_n$$

This follows immediately by the above theorem and using theorem (6 of section 3).

Remark: Strict inequalities may hold in the above theorem as can be seen by the following example.

Take $\{a_n\} = \{-1, 0, 1, -1, 0, 1, ...\}$

and $\{b_n\} = \{0, 1, -1, 0, 1, -1, ...\}$

Then,

$$-2 = \underline{\lim}\, a_n + \underline{\lim}\, b_n < -1 = \underline{\lim}\,(a_n + b_n) < 0$$

$$= \underline{\lim}\, a_n + \overline{\lim}\, b_n < 1 = \overline{\lim}\,(a_n + b_n) < 2 = \overline{\lim}\, a_n + \overline{\lim}\, b_n.$$

Theorem 18. *If $\{a_n\}$ and $\{b_n\}$ are two bounded sequences of non-negative real numbers, then*

(*i*) $(\underline{\lim}\, a_n)(\underline{\lim}\, b_n) \le \underline{\lim}\,(a_n b_n)$

(*ii*) $\underline{\lim}\,(a_n b_n) \le (\underline{\lim}\, a_n)(\overline{\lim}\, b_n)$

(*iii*) $(\underline{\lim}\, a_n)(\overline{\lim}\, b_n) \le \overline{\lim}\,(a_n b_n)$

(*iv*) $\overline{\lim}\,(a_n b_n) \le (\overline{\lim}\, a_n)(\overline{\lim}\, b_n)$.

We shall prove (*i*) and (*iii*) only, and leave the proof of (*ii*) and (*iv*) to the reader.

(*i*) We note that if $\underline{\lim}\, a_n = 0$ or $\underline{\lim}\, b_n = 0$, then the inequality (*i*) follows immediately. Therefore, we assume that

$$\underline{\lim}\, a_n = \underline{a} > 0 \text{ and } \underline{\lim}\, b_n = \underline{b} > 0,$$

Let $\varepsilon > 0$ be given. Then, by Theorem 8 (*ii*), there exist positive integers m_1 and m_2, respectively, such that

$$a_n > \underline{a} - \frac{\varepsilon}{2\underline{b}}, \ \forall\, n \ge m_1$$

and $$b_n > \underline{b} - \frac{\varepsilon}{2\underline{a}} \ \forall\, n \ge m_2.$$

CHAPTER 3

Therefore, for all $n \geq \max(m_1, m_2)$, we have

$$a_n b_n > \left(\underline{a} - \frac{\varepsilon}{2\underline{b}}\right)\left(\underline{b} - \frac{\varepsilon}{2\underline{a}}\right) = \underline{a}\,\underline{b} - \varepsilon + \frac{\varepsilon^2}{4\underline{a}\,\underline{b}} > \underline{a}\,\underline{b} - \varepsilon$$

This implies that, $\underline{\lim}(a_n b_n) \geq \underline{a}\,\underline{b} - \varepsilon$

But $\varepsilon > 0$ was arbitrary, hence the result follows.

$$\underline{\lim}(a_n b_n) \geq (\underline{\lim}\, a_n)(\underline{\lim}\, b_n)$$

(*iii*) We again suppose that

$$\underline{\lim}\, a_n = \underline{a} > 0, \text{ and } \overline{\lim}\, b_n = \bar{b} > 0$$

Then, (by Theorems 3.8 (*ii*) and 3.9 (*i*) of Section 3), for given $\varepsilon > 0$, there exists a positive integer m such that

$$a_n > \underline{a} - \frac{\varepsilon}{2\bar{b}}, \text{ for all } n \geq m,$$

and
$$b_n > \bar{b} - \frac{\varepsilon}{2\underline{a}}, \text{ for infinitely many values of } n.$$

Therefore, we have

$a_n b_n > \underline{a}\bar{b} - \varepsilon$, for infinitely many values of n.

Thus, using corollary 2 of Theorem 3.9, it follows that

$$\overline{\lim}(a_n b_n) \geq \underline{a}\bar{b} - \varepsilon$$

Hence,

$$\overline{\lim}(a_n b_n) \geq (\underline{\lim}\, a_n)(\overline{\lim}\, b_n),$$

since $\varepsilon > 0$ was arbitrary.

Remark: Strict inequalities may hold here as well.

Take

$$\{a_n\} = \{1, 2, 2, 1, 2, 2, ...\}$$

and

$$\{b_n\} = \{3, 1, 2, 3, 1, 2, ...\}$$

$\therefore$
$$\{a_n b_n\} = \{3, 2, 4, 3, 2, 4 ...\}$$

Then,

$$1 = (\underline{\lim}\, a_n)(\underline{\lim}\, b_n) < 2 = \underline{\lim}\, a_n b_n < 3 = (\underline{\lim}\, a_n)$$

$$(\overline{\lim}\, b_n) < 4 = \overline{\lim}(a_n b_n) < 6 = (\overline{\lim}\, a_n)(\overline{\lim}\, b_n).$$

Theorem 19. *If $\{a_n\}$ is a bounded sequence such that $a_n > 0$ for all $n \in \mathbf{N}$, then*

$$(i)\quad \underline{\lim}\left(\frac{1}{a_n}\right) = \frac{1}{\overline{\lim}\, a_n}, \text{ if } \overline{\lim}\, a_n > 0$$

and $$(ii)\quad \overline{\lim}\left(\frac{1}{a_n}\right) = \frac{1}{\underline{\lim}\, a_n}, \text{ if } \underline{\lim}\, a_n > 0.$$

(*i*) Since $a_n > 0$, $\forall\, n \in \mathbf{N}$, therefore

$$1 = \frac{1}{a_n} \cdot a_n, \quad \forall\, n \in \mathbf{N}$$

By using the inequality (*iii*) of the above theorem, we have

$$1 = \overline{\lim}\left(\frac{1}{a_n} a_n\right) \geq \left(\underline{\lim}\left(\frac{1}{a_n}\right)\right)\left(\overline{\lim}\, a_n\right)$$

i.e., $$\underline{\lim}\frac{1}{a_n} \leq \frac{1}{\overline{\lim}\, a_n} \qquad \left(\because\ \overline{\lim}\, a_n > 0\right) \qquad \text{...(1)}$$

Similarly, using the inequality (*ii*) of the above theorem, we obtain

$$1 = \underline{\lim}\left(\frac{1}{a_n} \cdot a_n\right) \leq \left(\underline{\lim}\frac{1}{a_n}\right)\left(\overline{\lim}\, a_n\right)$$

i.e., $$\underline{\lim}\frac{1}{a_n} \geq \frac{1}{\overline{\lim}\, a_n} \qquad \text{...(2)}$$

Hence from (1) and (2), we get

$$\underline{\lim}\frac{1}{a_n} = \frac{1}{\overline{\lim}\, a_n}, \text{ if } \overline{\lim}\, a_n > 0.$$

The proof of (*ii*) is similar.

Corollary. If $\{a_n\}$ and $\{b_n\}$ are bounded sequences, $a_n \geq 0$, $b_n > 0$ for all $n \in \mathbf{N}$, then

$$(i)\quad \underline{\lim}\left(\frac{a_n}{b_n}\right) \geq \frac{\underline{\lim}\, a_n}{\overline{\lim}\, b_n}, \text{ if } \overline{\lim}\, b_n > 0,$$

and $$(ii)\quad \overline{\lim}\left(\frac{a_n}{b_n}\right) \leq \frac{\overline{\lim}\, a_n}{\underline{\lim}\, b_n}, \text{ if } \underline{\lim}\, b_n > 0.$$

Follows directly from the above two theorems.

EXERCISE

1. Construct $\{a_n\}$, $\{b_n\}$ such that $a_{n+3} = a_n$, $b_{n+3} = b_n$, $a_n \neq 0$, $b_n \neq 0$

CHAPTER 3

$\underline{\lim}\, a_n \neq 0$, $\{a_n b_n\}$ and $\{b_n\}$ do not converge but

$$\underline{\lim}(a_n b_n) = \underline{\lim}\, a_n.\ \overline{\lim}\, b_n = \overline{\lim}\, a_n\ \underline{\lim}\, b_n$$

[*Hint*: $a_{3n} = 1$, $a_{3n-1} = 2$, $a_{3n-2} = -2$, $n = 1, 2, \ldots$
$b_{3n} = 2$, $b_{3n-1} = -2$, $b_{3n-2} = 1$, $n = 1, 2 \ldots$]

2. Construct $\{a_n\}$ such that $\overline{\lim}\, a_n > \underline{\lim}\ a_n > 0$

and either (*a*) $\overline{\lim}\left(a_n\, a_{n+1}\right) = \underline{\lim}\left(a_n\, a_{n+1}\right)$,

or (*b*) $\overline{\lim}\left(a_n\, a_{n+1}\, a_{n+2}\right) = \underline{\lim}\left(a_n\, a_{n+1}\, a_{n+2}\right)$.

Show that such a sequence cannot satisfy both (*a*) and (*b*).

[*Hint*: (*a*) Take $a_{2n} = 1$, $a_{2n-1} = 2$, $n = 1, 2, \ldots$
(*b*) Take $a_{3n} = 1$, $a_{3n-1} = 2$, $a_{3n-2} = 2$, $n = 1, 2 \ldots$]

3. The real bounded sequence $\{a_n\}$ is such that

$|a_n| \to l$, as $n \to \infty$, where l is a real number and

$\overline{\lim}\, a_n \neq \underline{\lim}\, a_n$. Show that $l \neq 0$ and $\overline{\lim}\, a_n = -\underline{\lim}\, a_n$

[*Hint*: If $|a_n| \to 0$, then $a_n \to 0$ and thus $\overline{\lim}\, a_n = \underline{\lim}\, a_n$.
Hence $l \neq 0$, show that $\overline{\lim}\, a_n = |l|$ and $\underline{\lim}\, a_n = -|l|$]

4. Let $\{a_n\}$ be any bounded positive sequence and $\{b_n\}$ is a convergent positive sequence. Show that

$$\overline{\lim}\,(a_n\, b_n) = \lim b_n\ \overline{\lim}\, a_n.$$

and

$$\underline{\lim}\,(a_n b_n) = \lim b_n\ \underline{\lim}\, a_n.$$

5. Let $\{a_n\}$ be a real bounded sequence and $\{b_n\}$ is a real convergent sequence. Show that

$$\overline{\lim}\,(a_n + b_n) = \overline{\lim}\, a_n + \lim b_n \text{ and}$$

$$\underline{\lim}\,(a_n + b_n) = \underline{\lim}\, a_n + \lim b_n$$

6. Let $\{b_n\}$ be a real bounded sequence such that for any real bounded sequence $\{a_n\}$, $\overline{\lim}\,(a_n + b_n) = \overline{\lim}\, a_n + \overline{\lim}\, b_n$. Show that $\{b_n\}$ is convergent.

7. Let $\{b_n\}$ be a positive real bounded sequence such that for any positive real bounded sequence $\{a_n\}$, $\overline{\lim}\,(a_n\, b_n) = \left(\overline{\lim}\, a_n\right)\left(\overline{\lim}\, b_n\right)$. Show that $\{b_n\}$ is convergent.

8. Let $\{a_n\}$ be a bounded sequence and z, ω be any given numbers (a_n, z, ω may be complex), show that:

$$\overline{\lim}\,|a_n - z| \leq |z - \omega| + \overline{\lim}\,|a_n - \omega|$$

$$\underline{\lim}\,|a_n - z| \geq |z - \omega| - \overline{\lim}\,|a_n - \omega|$$

Example 7. Show that $\lim \dfrac{(3n+1)(n-2)}{n(n+3)} = 3$.

■ We know that the sequence $\left\{\dfrac{1}{n}\right\}$ converges to zero, *i.e.*, $\lim \dfrac{1}{n} = 0$.

Now
$$\lim\frac{(3n+1)(n-2)}{n(n+3)}=\lim\frac{\left(3+\frac{1}{n}\right)\left(1-\frac{2}{n}\right)}{\left(1+\frac{3}{n}\right)}$$
$$=\frac{\lim\left(3+\frac{1}{n}\right)\lim\left(1-\frac{2}{n}\right)}{\lim\left(1+\frac{3}{n}\right)}=3.$$

EXERCISE

1. Show that

(i) $\lim\frac{3+2\sqrt{n+1}}{\sqrt{n+1}}=2$

(ii) $\lim\frac{2n^2-5}{3n^2+7n}=\frac{2}{3}$

(iii) $\lim\frac{1+2+3+...+n}{n^2}=\frac{1}{2}$

(iv) $\lim\frac{1+3+3+...+(2n-1)}{n^2}=1$

(v) $\lim[\sqrt{n+1}-\sqrt{n}]=0.$

2. Show that $\lim a_n=a\Rightarrow\lim|a_n|=|a|.$

Also by considering $a_n=(-1)^n$ or $(-1)^n\left(1+\frac{1}{n}\right)$,

show that the converse is not always true.
Is the converse true if $a=0$?

3. If $\{a_n\}$ converges and $\{b_n\}$ diverges, show that

(i) $\lim\frac{a_n}{b_n}=0.$

(ii) $\{a_n+b_n\}$ is divergent.

4. Given that $\lim a_n=a$, $\lim b_n=b$, and $\{S_n\}$ and $\{T_n\}$ are two sequences, where
$$S_n=\max(a_n, b_n)$$
$$T_n=\min(a_n, b_n).$$
Show that the sequences $\{S_n\}$ and $\{T_n\}$ are convergent and that
$$\lim S_n=\max(a, b),$$
$$\lim T_n=\min(a, b).$$

CHAPTER 3

[*Hint*: $\max(a_n, b_n) = \frac{1}{2}(a_n + b_n) + \frac{1}{2}|a_n - b_n|$

$\min(a_n, b_n) = \frac{1}{2}(a_n + b_n) - \frac{1}{2}|a_n - b_n|$.]

5. Use Cauchy's General Principle of Convergence to show that the following sequences are convergent:

(*i*) $\left\{\frac{n}{n+1}\right\}$,

(*ii*) $\left\{\frac{(-1)^n}{n}\right\}$,

(*iii*) $\left\{1 + \frac{1}{2!} + \frac{1}{3!} + \ldots + \frac{1}{n!}\right\}$,

(*iv*) $\left\{1 - \frac{1}{2} + \frac{1}{3} - \ldots + \frac{(-1)^{n-1}}{n}\right\}$.

8. SOME IMPORTANT THEOREMS

Theorem 20. *If* $\lim a_n = a$ *and* $a_n \geq 0$, *for all* n, *then* $a \geq 0$.

Let, if possible, $a < 0$.

Since $\lim a_n = a$, therefore for a given $\varepsilon > 0$, $\exists$ a positive integer m, such that

$$|a_n - a| < \varepsilon, \ \forall\, n \geq m$$

$$\Rightarrow \qquad a - \varepsilon < a_n < a + \varepsilon, \ \forall\, n \geq m$$

Let us select $\varepsilon = -\frac{1}{2}a, \lim_{n\to\infty}$ so that corresponding to $\varepsilon = -\frac{1}{2}a > 0$, $\exists$ a positive integer m_1 such that

$$a + \frac{1}{2}a < a_n < a - \frac{1}{2}a, \ \forall\, n \geq m_1$$

$$\Rightarrow \qquad \frac{3}{2}a < a_n < \frac{a}{2}, \ \forall\, n \geq m_1$$

or

$$a_n < \frac{a}{2}, \ \forall\, n \geq m_1$$

Since $a < 0$, therefore it follows that $a_n < 0$, for $n \geq m_1$. This contradicts the fact that $a_n > 0$, for all n. Therefore, the supposition is wrong.

Hence

$$a \nless 0 \Rightarrow a \geq 0.$$

Theorem 21. *If* $\{a_n\}, \{b_n\}$ *are two sequences such that*

(*i*) $a_n \leq b_n$, $\forall\, n$ *and*

(*ii*) $\lim a_n = a$, $\lim b_n = b$, *then* $a \leq b$.

Let, if possible, $a > b$.

Let $a - b = 3\varepsilon$, so that the neighbourhoods $]b-\varepsilon, b+\varepsilon[$, $]a-\varepsilon, a+\varepsilon[$ of b and a, respectively, are disjoint.

Since $\{a_n\}$, $\{b_n\}$ converge to a and b, respectively, therefore corresponding to $\varepsilon > 0$, $\exists$ positive integers m_1 and m_2, respectively, such that

$$a - \varepsilon < a_n < a + \varepsilon, \ \forall \ n \geq m_1$$
$$b - \varepsilon < b_n < b + \varepsilon, \ \forall \ n \geq m_2.$$

Let $\quad m = \max(m_1, m_2)$

$\therefore$
$$a_n \in \,]a-\varepsilon, a+\varepsilon[, \ \forall \ n \geq m$$
$$b_n \in \,]b-\varepsilon, b+\varepsilon[, \ \forall \ n \geq m$$

Consequently

$$b_n < a_n, \ \forall \ n \geq m$$

which contradicts the fact that

$$a_n \leq b_n, \ \forall \ n.$$

Hence, our supposition is wrong and therefore $a \leq b$.

Ex. 1. Deduce theorem 21 from theorem 20 by considering the sequence $\{c_n\}$, where $c_n = b_n - a_n$.

Ex. 2. If $\{a_n\}$, $\{b_n\}$ are two sequences, such that $a_n \leq b_n$, $\forall \ n$, then show that $\underline{\lim}\, a_n \leq \underline{\lim}\, b_n$, and $\overline{\lim}\, a_n \leq \overline{\lim}\, b_n$.

Theorem 22. ***Sandwich theorem.*** *If* $\{a_n\}$, $\{b_n\}$ and $\{c_n\}$ *are three sequences such that*

(*i*) $a_n \leq b_n \leq c_n, \ \forall \ n,$ ...(1)

and (*ii*) $\lim a_n = \lim c_n = l$

then
$$\lim b_n = l.$$

Let $\varepsilon > 0$ be given.

Now since $\{a_n\}$, $\{c_n\}$ both converge to l, therefore $\exists$ positive integers m_1, m_2, such that

$$|a_n - l| < \varepsilon, \ \forall \ n \geq m_1 \quad ...(2)$$

and
$$|c_n - l| < \varepsilon, \ \forall \ n \geq m_2 \quad ...(3)$$

Let $m = \max(m_1, m_2)$.

then, for $n \geq m$, we have from (1), (2) and (3)

$$l - \varepsilon < a_n \leq b_n \leq c_n < l + \varepsilon$$

$\Rightarrow$
$$l - \varepsilon < b_n < l + \varepsilon, \ \forall \ n \geq m$$

$\Rightarrow$
$$|b_n - l| < \varepsilon, \ \forall \ n \geq m.$$

Hence

$$\lim b_n = l.$$

Ex. If $\{a_n\}$, $\{b_n\}$ and $\{c_n\}$ are three sequences such that

$$a_n \le b_n \le c_n \quad \forall\, n$$

Then

(*i*) $\overline{\lim}\, a_n = \overline{\lim}\, c_n = \bar{l}$ implies $\overline{\lim}\, b_n = \bar{l}$, and

(*ii*) $\underline{\lim}\, a_n = \underline{\lim}\, c_n = \underline{l}$ implies $\underline{\lim}\, b_n = \underline{l}$

Example 8. Show that the sequence $\{b_n\}$, where

$$b_n = \left[\frac{1}{(n+1)^2} + \frac{1}{(n+2)^2} + \ldots + \frac{1}{(2n)^2}\right],$$

converges to zero.

■ Evidently,

$$\frac{n}{(2n)^2} \le b_n \le \frac{n}{n^2}, \quad \forall\, n$$

$$\Rightarrow \qquad \frac{1}{4n} \le b_n \le \frac{1}{n}, \quad \forall\, n.$$

Now the sequences $\{a_n\}$, $\{c_n\}$, where $a_n = \dfrac{1}{4n}$ and $c_n = \dfrac{1}{n}$ are such that

(*i*) $a_n \le b_n \le c_n$, $\forall\, n$, and

(*ii*) $\lim a_n = \lim c_n = 0$

$\therefore\ \lim b_n = 0.$

Ex. 1. Show that the sequence $\{b_n\}$, where

$$b_n = \left\{\frac{1}{\sqrt{(n^2+1)}} + \frac{1}{\sqrt{(n^2+2)}} + \ldots + \frac{1}{\sqrt{(n^2+n)}}\right\} = \sum_{k=1}^{n} \frac{1}{\sqrt{n^2+k}}$$

converges to 1.

Ex. 2. Show that:

(*i*) $\displaystyle \lim \sum_{k=1}^{n} \frac{1}{n^2+k} = 0,$

(*ii*) $\displaystyle \lim \sum_{k=1}^{n} \frac{1}{\sqrt{n+k}} = \infty.$

Theorem 23. *Cauchy's first theorem on limits. If* $\lim_{n\to\infty} a_n = l$, then

$$\lim_{n\to\infty}\left(\frac{a_1 + a_2 + \ldots + a_n}{n}\right) = l.$$

Let $b_n = a_n - l$.

Now, since $\lim a_n = l$, therefore

$$\lim b_n = 0.$$

Also
$$\frac{a_1 + a_2 + \ldots + a_n}{n} = l + \frac{b_1 + b_2 + \ldots + b_n}{n},\ \forall\, n$$

so that we have to show that

$$\lim \frac{b_1 + b_2 + \ldots + b_n}{n} = 0, \text{ when } \lim b_n = 0.$$

Now, since $\{b_n\}$ is convergent, therefore it is bounded and hence $\exists$ a number $K > 0$ such that

$$|b_n| \le K,\ \forall\, n.$$

Let $\varepsilon > 0$ be given. Since $\{b_n\}$ converges to zero, therefore $\exists$ a positive m such that

$$|b_n| < \tfrac{1}{2}\varepsilon, \text{ for } n \ge m$$

Also,

$$\begin{aligned}
\left|\frac{b_1 + b_2 + \ldots + b_n}{n}\right| &= \left|\frac{b_1 + b_2 + \ldots + b_m}{n} + \frac{b_{m+1} + \ldots + b_m}{n}\right| \\
&\le \frac{|b_1| + |b_2| + \ldots + |b_m|}{n} + \frac{|b_{m+1}| + \ldots + |b_n|}{n} \\
&< \frac{mK}{n} + \frac{\varepsilon}{2}\frac{(n-m)}{n},\ \forall\, n \ge m \\
&< \frac{mK}{n} + \frac{\varepsilon}{2}.
\end{aligned}$$

Let m_1 be a positive integer greater than $\dfrac{2mK}{\varepsilon}$, so that

$$\frac{mK}{n} < \frac{\varepsilon}{2}, \text{ where } n \ge m_1.$$

Thus, for $n \ge \max(m, m_1)$, we have

$$\left|\frac{b_1 + b_2 + \ldots + b_n}{n}\right| < \varepsilon$$

$\Rightarrow$
$$\lim \frac{b_1 + b_2 + \ldots + b_n}{n} = 0.$$

Hence,
$$\lim \frac{a_1 + a_2 + \ldots + a_n}{n} = l.$$

CHAPTER 3

Ex. Let $\{a_n\}$ be a bounded sequence.

Prove that

$$\underline{\lim}\, a_n \le \underline{\lim}\left((a_1 + a_2 + ... + a_n)/n\right) \le \overline{\lim}\left((a_1 + a_2 + ... + a_n)/n\right) \le \overline{\lim}\, a_n .$$

[*Hint*: For $\varepsilon > 0\ \exists m$ such that $a_n > \underline{\lim}\, a_n - \varepsilon,\ \forall\, n > m$

Thus

$$\frac{a_1 + a_2 + ... + a_n}{n} > (\underline{\lim}\, a_n - \varepsilon)\frac{n-m}{n} + \frac{a_1 + a_2 + ... + a_m}{n}, \text{ for } n > m$$

and

$$\underline{\lim} \frac{(a_1 + a_2 + ...\ a_n)}{n} \ge (\underline{\lim}\, a_n - \varepsilon) \lim \frac{n-m}{n} = \underline{\lim}\, a_n - \varepsilon .$$

This is true for all $\varepsilon > 0$. Hence the result].

Note: The converse of the theorem is not true.

Let $a_n = (-1)^n$, so that

$$\frac{a_1 + a_2 + ... + a_n}{n} = 0, \text{ if } n \text{ is even,}$$

$$= -\frac{1}{n}, \text{ if } n \text{ is odd.}$$

$$\Rightarrow \quad \lim \frac{a_1 + a_2 + ... + a_n}{n} = 0.$$

But $\{a_n\}$ is not convergent.

Example 9. Show that

$$\lim_{n\to\infty}\left[\frac{1}{\sqrt{n^2+1}} + \frac{1}{\sqrt{n^2+2}} + ... + \frac{1}{\sqrt{n^2+n}}\right] = 1.$$

■ Let

$$a_k = \frac{n}{\sqrt{n^2 + k}}, \ k = 1, 2 ... n$$

$$\therefore \qquad \lim a_n = \lim \frac{1}{\sqrt{1 + \frac{1}{n}}} = 1.$$

Thus by *Cauchy's first theorem on limits*, we have

$$\lim \frac{a_1 + a_2 + ... + a_n}{n} = 1.$$

or

$$\lim \frac{1}{n}\left[\frac{n}{\sqrt{n^2+1}}+\frac{n}{\sqrt{n^2+2}}+\ldots+\frac{n}{\sqrt{n^2+n}}\right]=1$$

or

$$\lim \left[\frac{1}{\sqrt{n^2+1}}+\frac{1}{\sqrt{n^2+2}}+\ldots+\frac{1}{\sqrt{n^2+n}}\right]=1.$$

Example 10. Show that

$$\lim_{n\to\infty} \frac{1}{n}\left[1+2^{1/2}+3^{1/3}+\ldots+n^{1/n}\right]=1.$$

- Consider $a_n = n^{1/n}$.
  Now $\lim a_n = \lim n^{1/n} = 1$
  $\therefore$ *By Cauchy's first theorem on limits,*

$$\lim \frac{a_1+a_2+\ldots+a_n}{n}=1$$

$$\Rightarrow \qquad \lim \frac{1}{n}\left[1+2^{1/2}+3^{1/3}+\ldots+n^{1/n}\right]=1.$$

Ex. Show that:

(*i*) $\lim \left[\frac{1}{\sqrt{n}}+\frac{1}{\sqrt{n+1}}+\ldots+\frac{1}{\sqrt{2n}}\right]=\infty$

(*ii*) $\lim \frac{1}{n}\left[1+\frac{1}{2}+\ldots+\frac{1}{n}\right]=0$

(*iii*) $\lim \left[\frac{1}{n^2}+\frac{1}{(n+1)^2}+\ldots+\frac{1}{(2n)^2}\right]=0$

Theorem 24. *If a sequence $\{a_n\}$ of positive terms converges to a positive limit l, then so does the sequences $\{(a_1a_2 \ldots a_n)^{1/n}\}$ of its geometric means, i.e., if $\lim_{n\to\infty} a_n = l$, then*

$$\lim_{n\to\infty} (a_1 a_2 \ldots a_n)^{1/n} = l.$$

Let $\lim_{n\to\infty} a_n = l$.

Since all terms are positive, the sequence $\{\log a_n\}$ of logarithms converges to $\log l$, *i.e.*, $\lim_{n\to\infty} \log a_n = \log l$.

CHAPTER 3

Hence by *Cauchy's first theorem on limits,*

$$\lim \frac{\log a_1 + \log a_2 + \dots + \log a_n}{n} = \log l$$

$$\Rightarrow \qquad \lim \log\left(a_1 a_2 \dots a_n\right)^{1/n} = \log l$$

$$\Rightarrow \qquad \lim\left(a_1\, a_2 \dots a_n\right)^{1/n} = l.$$

Note: The converse, however, is not true.

ILLUSTRATIONS

1. $$\left(1 \cdot \frac{2}{1} \cdot \frac{3}{2} \cdot \frac{4}{3} \cdots \frac{n}{n-1}\right)^{1/n} \to 1 \text{ as } n \to \infty$$

because $\lim \dfrac{n}{n-1} = 1$.

But $\left(1 \cdot \dfrac{2}{1} \cdot \dfrac{3}{2} \cdot \dfrac{4}{3} \cdots \dfrac{n}{n-1}\right)^{1/n} = (n)^{1/n}$

$$\therefore \qquad \lim_{n\to\infty} n^{1/n} = 1.$$

2. We know

$$\lim_{n\to\infty}\left(1 + \frac{1}{n}\right)^n = e$$

or

$$\lim_{n\to\infty}\left(\frac{n+1}{n}\right)^n = e$$

Hence, by theorem 3.24, as $n \to \infty$

$$\left\{\left(\frac{2}{1}\right)^1\left(\frac{3}{2}\right)^2\left(\frac{4}{3}\right)^3 \cdots \left(\frac{n+1}{n}\right)^n\right\}^{1/n} \to e$$

or

$$\frac{n+1}{(n!)^{1/n}} \to e$$

or

$$\frac{(n!)^{1/n}}{n+1} \to \frac{1}{e}$$

Note: $(n!)^{1/n} \to \infty$ as $n \to \infty$.

3. Let a sequence of positive monotonic decreasing terms

$$u_1, \frac{u_2}{u_1}, \frac{u_3}{u_2}, \dots, \frac{u_{n+1}}{u_n}, \dots \text{ converge to } l,$$

i.e.,
$$\lim_{n\to\infty} \frac{u_{n+1}}{u_n} = l.$$

Hence by theorem 3.24, the sequence

$$\left\{\left(u_1 \cdot \frac{u_2}{u_1} \cdot \frac{u_3}{u_2} \cdots \frac{u_{n+1}}{u_n}\right)^{1/n+1}\right\} \text{ converges to } l,$$

i.e.,
$$\left\{(u_{n+1})^{1/n+1}\right\} \text{ converges to } l$$

or

$$\lim_{n\to\infty} (u_n)^{1/n} = l$$

Thus,

$$\lim_{n\to\infty} \frac{u_{n+1}}{u_n} = l \;\Rightarrow\; \lim_{n\to\infty} (u_n)^{1/n} = l$$

The converse is not true.

A result that will be used while comparing the relative strength of the Ratio and the Root test for positive term infinite series. Also see Theorem 3.21.

Theorem 25. ***Cesaro's theorem.*** *If the sequences $\{a_n\}$ and $\{b_n\}$ converge to finite limits a and b respectively, then*

$$\lim_{n\to\infty} \frac{a_1 b_n + a_2 b_{n-1} + \dots + a_n b_1}{n} = ab$$

Let $a_n = a + \alpha_n$, where $|\alpha_n| \to 0$ as $n \to \infty$. On substituting for $a_1, a_2, \dots, a_n$, we get

$$\frac{a_1 b_n + a_2 b_{n-1} + \dots a_n b_1}{n} = \frac{a(b_1 + b_2 + \dots + b_n)}{n} + \frac{\alpha_1 b_n + \alpha_2 b_{n-1} + \dots + \alpha_n b_1}{n}$$

$$\le \frac{a(b_1 + b_2 + \dots + b_n)}{n} + \frac{B(|\alpha_1| + |\alpha_2| + \dots + |\alpha_n|)}{n}$$

where B is the upper bound of the numbers $|b_1|, |b_2|, \dots$

Also by *Cauchy's first theorem on limits*

$$\lim \frac{b_1 + b_2 + \dots + b_n}{n} = b$$

CHAPTER 3

and

$$\lim \frac{|\alpha_1| + |\alpha_2| + ... |\alpha_n|}{n} = 0$$

$$\therefore \qquad \lim \frac{a_1 b_n + a_2 b_{n-1} + ... a_n b_1}{n} = ab$$

Theorem 26. ***Cauchy's second theorem on limits.*** *If all the terms of a sequence $\{u_n\}$ are positive and if $\lim\limits_{n\to\infty} \frac{u_{n+1}}{u_n}$ exists then so does $\lim\limits_{n\to\infty} (u_n)^{1/n}$, and the two limits are equal, i.e., $\lim (u_n)^{1/n} = \lim \frac{u_{n+1}}{u_n}$, provided the latter limit exists.*

Let $\lim\limits_{n\to\infty} \frac{u_{n-1}}{u_n} = l$, a finite real number.

Hence for any $\varepsilon > 0$ there exists a positive integer m, such that

$$l - \varepsilon < \frac{u_{n+1}}{u_n} < l + \varepsilon, \ \forall\, n \ge m.$$

Putting $n = m, m + 1, ..., n - 1$ and multiplying, we get

$$(l - \varepsilon)^{n-m} < \frac{u_n}{u_m} < (l + \varepsilon)^{n-m}, \ \forall\, n \ge m$$

or

$$(l - \varepsilon)^{n-m} u_m < u_n < (l + \varepsilon)^{n-m} u_m, \ \forall\, n \ge m$$

or

$$(l - \varepsilon)^{1-m/n} u_m^{1/n} < u_n^{1/n} < (l + \varepsilon)^{1-m/n} u_m^{1/n}, \ \forall n \ge m.$$

If the first and the last expression in the above inequality are called A_n and B_n, then

$$\lim A_n = l - \varepsilon \text{ and } \lim B_n = l + \varepsilon$$

because u_m is a finite positive quantity.

It is, therefore, possible to choose a positive integer m_o such that for $n \ge m_o$,

$$A_n > l - 2\varepsilon \text{ and } B_n < l + 2\varepsilon$$

$$\therefore \qquad l - 2\varepsilon < u_n^{1/n} < l + 2\varepsilon, \ \forall\, n \ge \max(m, m_0)$$

i.e.,

$$\lim u_n^{1/n} = l$$

Corollary. More generally, if $\overline{\lim} \frac{u_{n+1}}{u_n} = U$ and $\underline{\lim} \frac{u_{n+1}}{u_n} = L$, we can modify the inequalities by replacing $l - \varepsilon$ by $L - \varepsilon$ and $l + \varepsilon$ by $U + \varepsilon$. Thus

$$L - 2\varepsilon < u_n^{1/n} < U + 2\varepsilon, \ \forall\, n \ge \max(m, m_0)$$

$$\Rightarrow \qquad L \le \underline{\lim}\, u_n^{1/n} \text{ and } \overline{\lim}\, u_n^{1/n} \le U$$

$$\therefore \qquad \underline{\lim}\frac{u_{n+1}}{u_n} \le \underline{\lim}\, u_n^{1/n} \le \overline{\lim}\, u_n^{1/n} \le \overline{\lim}\frac{u_{n+1}}{u_n},$$

a very useful result.

Note: Taking $u_{2n-1} = \dfrac{1}{2^n}$ and $u_{2n} = \dfrac{1}{3^n}$, for $n = 1, 2, 3, \ldots$

$$0 = \underline{\lim}\frac{u_{n+1}}{u_n} < \frac{1}{\sqrt{3}} = \underline{\lim}\, u_n^{1/n} < \frac{1}{\sqrt{2}} = \overline{\lim}\, u_n^{1/n} < \infty = \overline{\lim}\frac{u_{n+1}}{u_n}.$$

Example 11. Show that the sequences $\{a_n^{1/n}\}$ and $\{b_n^{1/n}\}$, where

(i) $a_n = \dfrac{(3n)!}{(n!)^3}$ and

(ii) $b_n = \dfrac{n^n}{(n+1)(n+2)\ldots(n+n)}$

converge and find their limits.

■ *(i)*

$$a_n = \frac{(3n)!}{(n!)^3}$$

$$a_{n+1} = \frac{(3n+3)!}{((n+1)!)^3}$$

$$\lim\frac{a_{n+1}}{a_n} = \lim\frac{(3n+3)!}{((n+1)!)^3}\,\frac{(n!)^3}{(3n)!}$$

$$= \lim\frac{(3n+3)(3n+2)(3n+1)}{(n+1)^3} = 27$$

By *Cauchy's second theorem on limits,*

$$\lim a_n^{1/n} = \lim\frac{a_{n+1}}{a_n} = 27$$

(ii) $b_n = \dfrac{n^n}{(n+1)(n+2)\ldots(n+n)}$

$$b_{n+1} = \frac{(n+1)^{n+1}}{(n+2)(n+3)\ldots 2n(2n+1)(2n+2)}$$

$$\therefore \qquad \lim \frac{b_{n+1}}{b_n} = \lim \frac{(n+1)^{n+1}(n+1)(n+2)\dots(2n)}{(n+2)(n+3)\dots(2n+1)(2n+2)n^n}$$

$$= \lim \frac{\left(1+\frac{1}{n}\right)^n \left(1+\frac{1}{n}\right)^2}{\left(2+\frac{1}{n}\right)\left(2+\frac{2}{n}\right)} = \frac{e}{4}$$

Hence by *Cauchy's second theorem on limits,*

$$\lim b_n^{1/n} = \lim \frac{b_{n+1}}{b_n} = \frac{e}{4}$$

Ex. Find $\lim\limits_{n\to\infty} \frac{1}{n}\left((m+1)(m+2)\dots(m+n)\right)^{1/n}$.

Theorem 27. *If* $\{a_n\}$ *be a sequence such that* $\lim \frac{a_{n+1}}{a_n} = l$, *where* $|l| < 1$, *then* $\lim a_n = 0$.

Since $|l| < 1$, we can choose a positive number ε, so small such that $|l| + \varepsilon < 1$.

Now since $\lim \frac{a_{n+1}}{a_n} = l$, therefore $\exists$ a positive integer m, such that

$$\left|\frac{a_{n+1}}{a_n} - l\right| < \varepsilon, \forall n \geq m$$

$$\Rightarrow \qquad \left|\frac{a_{n+1}}{a_n}\right| - |l| \leq \left|\frac{a_{n+1}}{a_n} - l\right| < \varepsilon, \forall n \geq m$$

$$\Rightarrow \qquad \left|\frac{a_{n+1}}{a_n}\right| < |l| + \varepsilon = k \text{ (say), where } k < 1, \ \forall n \geq m$$

Putting $n = m, m+1, \dots, (n-1)$ in turn and multiplying, we get

$$\left|\frac{a_n}{a_m}\right| < k^{n-m}, \ \forall n \geq m$$

$$\Rightarrow \qquad |a_n| < \frac{|a_m|}{k^m} \cdot k^n, \forall n \geq m$$

But as $k < 1$, $k^n \to 0$, hence

$$\lim a_n = 0.$$

Theorem 28. *If* $\{a_n\}$ *be a sequence, such that* $\lim \frac{a_{n+1}}{a_n} = l > 1$, *then*

$$\lim a_n = \infty.$$

Since $l > 1$, we can choose a positive number ε, such that

$$l - \varepsilon > 1.$$

Now, since $\lim \dfrac{a_{n+1}}{a_n} = l$, therefore $\exists$ a positive integer m, such that

$$\left| \frac{a_{n+1}}{a_n} - l \right| < \varepsilon, \ \forall\, n \geq m$$

$$\Rightarrow \qquad l - \varepsilon < \frac{a_{n+1}}{a_n} < l + \varepsilon, \ \forall\, n \geq m$$

$$\Rightarrow \qquad \frac{a_{n+1}}{a_n} > l - \varepsilon = k \text{ (say)}, \ \forall\, n \geq m \text{ where } k > 1$$

Putting $n = m, m + 1, \ldots, (n - 1)$ in turn and multiplying, we get

$$\left| \frac{a_n}{a_m} \right| \geq \frac{a_n}{a_m} > k^{n-m}, \ \forall\, n \geq m$$

$$\Rightarrow \qquad |a_n| > \frac{|a_m|}{k^m} k^n, \ \forall\, n \geq m$$

But as $k > 1$, $k^n \to \infty$

$$\therefore \qquad \lim a_n = \infty$$

Example 12. Show that for any real number x, $\lim \dfrac{x^n}{n!} = 0$.

■ Let $a_n = \dfrac{x^n}{n!}$

$$\therefore \qquad \frac{a_{n+1}}{a_n} = \frac{x^{n+1}}{(n+1)!} \frac{n!}{x^n} = \frac{x}{n+1},$$

so that
$$\lim \frac{a_{n+1}}{a_n} = 0 < 1.$$

Hence by theorem 27, $\lim a_n = 0$.

Example 13. Show that

$$\lim \frac{m(m-1)(m-2)\ldots(m-n+1)}{n!} x^n = 0, |x| < 1$$

■ Let

$$a_n = \frac{m(m-1)(m-2)\ldots(m-n+1)}{n!} x^n$$

$$\therefore \quad \lim \frac{a_{n+1}}{a_n} = \lim \left(\frac{m-n}{n+1} \right) x = \lim \frac{\frac{m}{n} - 1}{1 + \frac{1}{n}} x = -x$$

But $|-x| = |x| < 1$.

$\therefore$ By theorem 3.27, $\lim a_n = 0$.

Ex. Show that $\lim nr^n = 0$, if $|r| < 1$.

9. MONOTONIC SEQUENCES

A sequence $\{S_n\}$ is said to be *monotonic increasing*, if $S_{n+1} \geq S_n \ \forall \ n$, and *monotonic decreasing* if $S_{n+1} \leq S_n \ \forall \ n$. It is said to be monotonic if it is either monotonic increasing or monotonic decreasing.

A sequence (S_n) is *strictly increasing* if $\forall \ n, S_{n+1} > S_n$ and *strictly decreasing* if $S_{n+1} < S_n$.

The importance of monotonic sequences lies in the fact that they cannot oscillate. They either converge or diverge. Following theorem will make the point clear.

Theorem 29. *A necessary and sufficient condition for the convergence of a monotonic sequence is that it is bounded.*

The condition is necessary for we know that every convergent sequence is bounded.

The condition is sufficient. Let a bounded sequence $\{S_n\}$ be monotonic increasing. Let S denotes its range, which is evidently bounded. By the completeness property, S has the least upper bound (the supremum), say M.

We shall show that $\{S_n\}$ converges to M.

Let ε be any pre-assigned positive number.

Now since $M - \varepsilon$ is a number less than the supremum M, there exists at least one member say S_m such that $S_m > M - \varepsilon$.

As $\{S_n\}$ is a monotonic increasing sequence,

$$\therefore \quad S_n \geq S_m > M - \varepsilon, \quad \forall \ n \geq m$$

Again, since M is the supremum,

$$S_n \leq M < M + \varepsilon, \quad \forall \ n$$

Thus

$$M - \varepsilon < S_n < M + \varepsilon, \quad \forall \ n \geq m$$

$$\Rightarrow \quad |S_n - M| < \varepsilon, \quad \forall \ n \geq m$$

$$\Rightarrow \quad \{S_n\} \text{ converges and } \lim S_n = M.$$

We may similarly consider the case of a bounded monotonic decreasing sequence.

Corollary 1. A monotonic increasing bounded above sequence converges to its least upper bound and a monotonic decreasing bounded below to the greatest lower bound.

Corollary 2. Every monotonic increasing sequence which is not bounded above, diverges to $+\infty$.

Let $\{S_n\}$ be a monotonic increasing sequence, not bounded above. Let G be any real number however large.

Since the sequence $\{S_n\}$ is unbounded above the monotonic increasing, $\exists$ a positive integer m such that

$$S_m > G \text{ and } S_n \geq S_m, \quad \forall\, n \geq m$$

$$\Rightarrow \qquad S_n > G, \quad \forall\, n \geq m$$

Hence, $\lim S_n = +\infty$.

Corollary 3. Every monotonic decreasing sequence which is not bounded below, diverges to $-\infty$.

9.1 Subsequences

If $\{S_n\} = \{S_1, S_2, S_3, ...\}$ be a sequence, then any infinite succession of its terms, picked out in any way (but preserving the original order), is called a *subsequence* of $\{S_n\}$, or, in other words if $\{n_k\}$ be a strictly monotonic increasing sequence of natural numbers, *i.e.*, $n_1 < n_2 < n_3 < ...$, then $\{S_{n_k}\}$ is a *subsequence* of the sequence $\{S_n\}$.

ILLUSTRATIONS

1. $\{S_2, S_4, S_6, ..., S_{2n}, ...\}$ is a subsequence of $\{S_n\}$.
2. $\{S_1, S_4, S_9, ..., S_{n^2}, ...\}$ is a subsequence of $\{S_n\}$.
3. $\{S_7, S_8, S_9, ...\}$ is a subsequence of $\{S_n\}$, which is obtained by removing a finite number of terms from the beginning of $\{S_n\}$.

Without going into a formal proof we state :

1. *A sequence $\{S_n\}$ converges to s if and only if its every subsequence converges to s. Similarly* $\lim S_n = \infty(-\infty)$ *if and only if every subsequence of $\{S_n\}$ tends to* $\infty(-\infty)$.
2. *If ξ is a limit point of a sequence $\{S_n\}$, then there exists a subsequence $\{S_{n_k}\}$ of $\{S_n\}$ which converges to ξ, i.e.,* $\lim\limits_{k\to\infty} S_{n_k} = \xi$.

Ex. The subsequence $\{x, x^4, x^9, x^{16}, ..., x^{n^2}, ...\}$ of $\{x^n\}$ converges to zero if $|x| < 1$, for the sequence $\{x^n\}$ converges to zero for $|x| < 1$.

Example 14. Show that the sequence $\{S_n\}$, where $S_n = \left(1 + \frac{1}{n}\right)^n$, is convergent and that $\lim \left(1 + \frac{1}{n}\right)^n$ lies between 2 and 3.

■ Expanding by Binomial theorem, since n is a positive integer, we get

CHAPTER 3

$$S_n = \left(1+\frac{1}{n}\right)^n = 1 + n\frac{1}{n} + \frac{n(n-1)}{2!}\cdot\frac{1}{n^2} + \ldots + \frac{n(n-1)\ldots 3\cdot 2\cdot 1}{n!}\cdot\frac{1}{n^n}$$

$$= 1 + 1 + \frac{1}{2!}\left(1-\frac{1}{n}\right) + \frac{1}{3!}\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right) + \ldots + \frac{1}{n!}\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\ldots\left(1-\frac{n-1}{n}\right)$$

Since each term beyond the first two terms on the R.H.S. is an increasing function of n, it follows that $\{S_n\}$ is a monotonic increasing sequence. Again since each bracket on the R.H.S. is positive, therefore, we have

$$2 < S_n < 1 + \frac{1}{1} + \frac{1}{2!} + \frac{1}{3!} + \ldots + \frac{1}{n!}$$

$$< 1 + 1 + \frac{1}{2} + \frac{1}{2^2} + \ldots + \frac{1}{2^{n-1}} < 3$$

$\Rightarrow$ $$2 < S_n < 3$$

Thus $\{S_n\}$ is a bounded and monotonic increasing sequence and so has a limit, which is generally denoted by e.

$\therefore$ $$\lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n = e, \text{ where } 2 < e < 3$$

Example 15. Show that the sequence $\{S_n\}$, where

$$S_n = \frac{1}{n+1} + \frac{1}{n+2} + \ldots + \frac{1}{n+n}, \quad \forall\, n \in \mathbf{N}$$

is convergent.

■ Now

$$S_{n+1} - S_n = \frac{1}{2n+1} + \frac{1}{2n+2} - \frac{1}{n+1}$$

$$= \frac{1}{2(n+1)(2n+1)} > 0, \quad \forall\, n$$

$\therefore$ The sequence $\{S_n\}$ is monotonic increasing.

Again

$$S_n = \frac{1}{n+1} + \frac{1}{n+2} + \ldots + \frac{1}{n+n}$$

$$< \frac{1}{n} + \frac{1}{n} + \ldots + \frac{1}{n} = 1$$

i.e., $$0 < S_n < 1$$

$\therefore$ The sequence is bounded.

Hence, the sequence being bounded and monotonic increasing, is convergent.

Example 16. Show that the sequence, $\{S_n\}$, where

$$S_n = \frac{1}{1!} + \frac{1}{2!} + \dots + \frac{1}{n!}, \quad \forall\, n \in \mathbf{N}$$

is convergent.

■ Now

$$S_{n+1} - S_n = \frac{1}{n+1!} > 0, \forall\, n$$

$\therefore$ The sequence $\{S_n\}$ is monotonic increasing.

Again,

$$S_n = \frac{1}{1!} + \frac{1}{2!} + \dots + \frac{1}{n!}$$

$$< 1 + \frac{1}{2} + \frac{1}{2^2} + \dots + \frac{1}{2^{n-1}}$$

$$= \frac{1 - \frac{1}{2^n}}{1 - \frac{1}{2}} = 2 - \frac{1}{2^{n-1}} < 2$$

$\Rightarrow$ $\quad 0 < S_n < 2$

$\therefore$ The sequence is bounded.

Thus, the sequence $\{S_n\}$, being bounded and monotonic increasing, is convergent.

Example 17. Show that the sequence $\{S_n\}$, defined by the recursion formula $S_{n+1} = \sqrt{3S_n}$, $S_1 = 1$, converges to 3.

■ The terms of the sequence $\{S_n\}$ are

$$1, \sqrt{3}, \sqrt{3\sqrt{3}}, \sqrt{3\sqrt{3\sqrt{3}}}, \dots$$

Clearly

$$S_2 > S_1$$
$$S_3 > S_2$$

Also

$$S_{m+1} > S_m \Rightarrow \sqrt{3S_{m+1}} > \sqrt{3S_m} \Rightarrow S_{m+2} > S_{m+1}.$$

Thus by *Mathematical Induction*, the sequence $\{S_n\}$ is monotonic increasing.

Again,

$$S_1 < 3$$

$$S_2 < 3$$

$$S_3 = \sqrt{3S_2} < 3$$

and

$$S_m < 3 \Rightarrow \sqrt{3S_m} < \sqrt{3.3} \Rightarrow S_{m+1} < 3$$

So, again, by *Mathematical Induction*,

$$0 < S_n < 3, \ \forall \ n$$

Hence, the sequence $\{S_n\}$, being bounded and monotonic increasing, is convergent.
Let $\lim S_n = l$.

Since $\lim S_{n+1} = \lim \sqrt{3S_n}$

$\therefore$ $$l = \sqrt{3l} \Rightarrow l = 0 \text{ or } 3$$

But $l \neq 0$, since $S_n \geq 1, \ \forall \ n$

Hence, $\lim S_n = 3$.

Example 18. If $\{S_n\}$ is a sequence, such that

$$S_{n+1} = \sqrt{\frac{ab^2 + S_n^2}{a+1}}, \ b > a, \ \forall \ n \geq 1 \text{ and } S_1 = a > 0$$

then show that the sequence $\{S_n\}$ is an increasing bounded above sequence and $\lim\limits_{n\to\infty} S_n = b$.

■ Given that

$$S_{n+1} = \sqrt{\frac{ab^2 + S_n^2}{a+1}}, \ b > a, \ \forall \ n \geq 1$$

For $n = 1$

$$S_2 = \sqrt{\frac{ab^2 + S_1^2}{a+1}} = \sqrt{\frac{ab^2 + a^2}{a+1}},$$

and $$S_2^2 - S_1^2 = \frac{ab^2 + a^2}{a+1} - a^2 = \frac{a\left(b^2 - a^2\right)}{a+1} > 0$$

$\therefore$ $$S_2 > S_1 \qquad (\because b > a > 0)$$

For $n \geq 1$, we have

$$S_{n+1}^2 - S_n^2 = \frac{ab^2 + S_n^2}{a+1} - S_n^2 = \frac{a\left(b^2 - S_n^2\right)}{a+1} \qquad \ldots(1)$$

Now,

$$S_1 = a < b, b^2 - S_2^2 = \frac{b^2 - a^2}{a+1} > 0$$

$\therefore$ $$S_2^2 < b^2$$

Assume $S_m < b$, for some $m \in \mathbf{N}$, then

$$S_{m+1}^2 - b^2 = \frac{ab^2 + S_m^2}{a+1} - b^2 = \frac{S_m^2 - b^2}{a+1} < 0$$

So by the Principle of Mathematical Induction,

$$0 < S_n < b, \ \forall \ n \in \mathbf{N}$$

$\therefore$ The sequence $\{S_n\}$ is bounded.

Also (1) implies that the sequence $\{S_n\}$ is monotonic increasing. Hence, the sequence $\{S_n\}$, being bounded and monotonically increasing, is convergent.

Let $$\lim_{n\to\infty} S_n = l$$

Since $$\lim_{n\to\infty} S_{n+1} = \lim_{n\to\infty} \sqrt{\frac{ab^2 + S_n^2}{a+1}}$$

$\therefore$ $$l^2 = \frac{ab^2 + l^2}{a+1} \Rightarrow l = \pm b \qquad (\because \ a \neq 0)$$

But $l \neq -b$ as $S_n \geq a > 0, \ \forall \ n$

Hence $l = b$.

Example 19. If $\{S_n\}$ be a sequence of positive real numbers such that $S_n = \frac{1}{2}(S_{n-1} + S_{n-2}), \ \forall \ n > 2$, then show that $\{S_n\}$ converges. Also find $\lim S_n$.

■ If $S_1 = S_2$, then evidently $S_n = S_1, \ \forall \ n$

so that the sequence converges to S_1.

When $S_1 \neq S_2$, let $S_1 < S_2$.

Putting $n = 3, 4, 5, \ldots, m$, in the relation $S_n = \frac{1}{2}(S_{n-1} + S_{n-2})$, we find that

$$\left.\begin{aligned} S_3 &= \frac{1}{2}(S_2 + S_1) \\ S_4 &= \frac{1}{2}(S_3 + S_2) \\ S_5 &= \frac{1}{2}(S_4 + S_3) \\ &\cdots\cdots\cdots\cdots \\ S_m &= \frac{1}{2}(S_{m-1} + S_{m-2}) \\ &\cdots\cdots\cdots\cdots \end{aligned}\right\} \quad \ldots(1)$$

$\Rightarrow$
$$\begin{aligned} S_1 &< S_3 < S_2 \\ S_3 &< S_4 < S_2 \\ S_3 &< S_5 < S_4 \\ S_5 &< S_6 < S_4 \\ &\cdots\cdots\cdots \end{aligned}$$

CHAPTER 3

Thus it appears that

and
$$\left.\begin{aligned} S_1 < S_3 < S_5 < \ldots \\ S_2 > S_4 > S_6 > \ldots \end{aligned}\right\} \qquad \ldots(2)$$

Now

$$S_{n+2} - S_n = \frac{1}{2}\left(S_{n+1} + S_n\right) - S_n$$

$$= \frac{1}{2}\left(S_{n+1} - S_n\right) \qquad \ldots(3)$$

$$= \frac{1}{4}\left(S_n - S_{n-2}\right) \qquad \ldots(4)$$

From (4), we easily see that

(*i*) The subsequence of odd terms is monotonic increasing, *i.e.*,

$$S_1 < S_3 < S_5 < \ldots$$

(*ii*) The subsequence of even terms is monotonic decreasing, *i.e.*,

$$\ldots < S_6 < S_4 < S_2.$$

Again from (3) when n is even, putting $n = 2m$, we get

$$S_{2m+2} - S_{2m} = \frac{1}{2}\left(S_{2m+1} - S_{2m}\right)$$

$\therefore$
$$S_{2m+2} < S_{2m} \Rightarrow S_{2m+1} < S_{2m}$$

but

$$S_{2m} < S_{2m-2} < \ldots < S_4 < S_2$$

$\therefore$ Every odd term is less than every even term, *i.e.*,

$$S_1 < S_3 < S_5 < \ldots < S_{2m+1} < S_{2m} < S_{2m-2} < \ldots < S_6 < S_4 < S_2$$

Thus, the odd term subsequence $\{S_{2n+1}\}$ is monotonic increasing and is bounded above (by S_2) and is, therefore, convergent.

Similarly the even term subsequence $\{S_{2n}\}$ is convergent.

Let us now show that the two subsequences converge to the same number.

Let $\{S_{2n+1}\} \to l$ and $\{S_{2n}\} \to l'$.

From the recursion formula,

$$S_{2m} = \frac{1}{2}\left(S_{2m-1} + S_{2m-2}\right)$$

On taking limits as $m \to \infty$, we get

$$l' = \frac{1}{2}(l + l') \;\Rightarrow\; l = l'.$$

Thus both the subsequences of $\{S_n\}$ converge to l.

$\Rightarrow$ The sequence $\{S_n\}$ converges to l.

Again from (1) by adding,

$$S_k + \frac{1}{2}S_{k-1} = \frac{1}{2}(S_1 + 2S_2)$$

Taking limits when $k \to \infty$, we get

$$l + \frac{1}{2}l = \frac{1}{2}(S_1 + 2S_2)$$

$\therefore$
$$l = \frac{1}{3}(S_1 + 2S_2)$$

Thus the sequence $\{S_n\}$ converges to $\frac{1}{3}(S_1 + 2S_2)$.

It may similarly be shown that the case $S_1 > S_2$ leads to the same result.

Example 20. Show that the sequence $\{a_n\}$ defined by

$$a_{n+1} = \frac{1}{2}\left(a_n + \frac{9}{a_n}\right),\ n \geq 1 \text{ and } a_1 > 0$$

converges to 3.

■ Now
$$a_2 - a_1 = \frac{1}{2}\left(a_1 + \frac{9}{a_1}\right) - a_1 = \frac{9 - a_1^2}{2a_1} \geq 0, \text{ if } a_1 \leq 3$$

$\Rightarrow$ $\quad a_2 \geq a_1$, if $a_1 \leq 3$

Also
$$a_{n+1} - a_n = \frac{1}{2}\left(a_n + \frac{9}{a_n}\right) - a_n = \frac{9 - a_n^2}{2a_n} \geq 0, \text{ if } a_n \leq 3.$$

Thus the sequence $\{a_n\}$ is monotonically increasing if $a_n \leq 3,\ \forall\ n$ and decreasing if $a_n \geq 3,\ \forall\ n$. In either case the sequence is monotonic and bounded and therefore it is convergent.

let
$$\lim_{n\to\infty} a_n = l.$$

Now,
$$\lim_{n\to\infty} a_{n+1} = \lim_{n\to\infty} \frac{1}{2}\left(a_n + \frac{9}{a_n}\right) = \frac{1}{2}\left(l + \frac{9}{l}\right)$$

or $\quad l^2 = 9$

or $\quad l = \pm 3.$

But $\quad l \neq -3$ as $a_n > 0,\ \forall\ n.$

Hence $l = 3$.

Example 21. If a sequence $\{S_n\}$ is defined by

$$S_n = \frac{S}{1 + S_{n-1}},\ \text{where } S > 0, S_1 > 0,\ n \geq 2$$

then show that the sequence converges to the positive root of the equation

$$x^2 + x - S = 0.$$

- Now

$$S_n = \frac{S}{1 + S_{n-1}} \qquad \text{...(1)}$$

$$S_n - S_{n-2} = \frac{S}{1 + S_{n-1}} - \frac{S}{1 + S_{n-3}}$$

$$= \frac{-S\left(S_{n-1} - S_{n-3}\right)}{\left(1 + S_{n-1}\right)\left(1 + S_{n-3}\right)} \qquad \text{...(2)}$$

$$= \frac{S^2\left(S_{n-2} - S_{n-4}\right)}{\left(1 + S_{n-1}\right)\left(1 + S_{n-2}\right)\left(1 + S_{n-3}\right)\left(1 + S_{n-4}\right)} \qquad \text{...(3)}$$

(3) shows that the even and odd terms form separate monotonic subsequences.

Again (2) shows that if odd terms form a monotonic decreasing subsequence, even terms will form a monotonic increasing subsequence and vice-versa.

Since every term of the sequence is positive,

$$\therefore \qquad S - S_n = \frac{SS_{n-1}}{1 + S_{n-1}} > 0$$

$$\Rightarrow \qquad 0 < S_n < S, \ \forall\, n$$

a result which could have been written directly from (1).

Thus the monotonic increasing subsequence is bounded above by S and the monotonic decreasing subsequence is bounded below by 0.

Hence the two subsequences converge.

Let, if possible, the even term subsequence converge to l and the odd term subsequence to l'.

$\therefore$ Taking the limit, we get from (1)

(*i*) for n even, $l = \dfrac{S}{1 + l'}$ or $ll' + l = S$

(*ii*) for n odd, $l' = \dfrac{S}{1 + l}$ or $ll' + l' = S$

$$\Rightarrow \qquad l = l'$$

Thus both the subsequences converge to the same number l,

$\Rightarrow$ The given sequence $\{S_n\}$ converges to l.

Again from (1), on proceeding to limits we get

$$l = \frac{S}{1 + l} \Rightarrow l^2 + l - S = 0$$

$\therefore$ l is a root of $x^2 + x - S = 0$

where l is positive, for, every term of the sequence is positive.

Example 22. Two sequences $\{x_n\}$ and $\{y_n\}$ are defined inductively by

$$x_1 = \frac{1}{2} \text{ and } y_1 = 1$$

and

$$x_n = \sqrt{x_{n-1}\, y_{n-1}}, \qquad n = 2, 3, 4, \ldots$$

$$\frac{1}{y_n} = \frac{1}{2}\left(\frac{1}{x_n} + \frac{1}{y_{n-1}}\right), \qquad n = 2, 3, 4, \ldots$$

Prove that

$$x_{n-1} < x_n < y_n < y_{n-1}, \quad n = 2, 3, \ldots$$

and deduce that both the sequences converge to the same limit l, where $\frac{1}{2} < l < 1$.

- If $0 < a < b$, then geometric mean $G = \sqrt{ab}$ and the harmonic mean

$$H = \left[\frac{1}{2}\left(\frac{1}{a} + \frac{1}{b}\right)\right]^{-1}$$

Also

$$a < H < G < b$$

we are given that

$$\frac{1}{2} = x_1 < y_1 = 1$$

On the assumption $x_{n-1} < y_{n-1}$, we have

$$x_{n-1} < x_n < y_{n-1} \qquad \left(\because\ x_n = \sqrt{x_{n-1}\, y_{n-1}}\right)$$

Further

$$x_n < y_n < y_{n-1}$$

because y_n is the harmonic means of x_n and y_{n-1}. It follows by induction that

$$x_{n-1} < x_n < y_n < y_{n-1}, \quad n = 2, 3, \ldots$$

The sequence $\{x_n\}$ increases and is bounded above by $y_1 = 1$. The sequence $\{y_n\}$ decreases and is bounded by $x_1 = \frac{1}{2}$. Hence, both sequences converge. Suppose $x_n \to l$ as $n \to \infty$ and $y_n \to m$ as $n \to \infty$, then

$$l^2 = lm$$

and

$$\frac{1}{m} = \frac{1}{2}\left(\frac{l+m}{lm}\right)$$

Both the sequences yield $l = m$.

CHAPTER 3

Theorem 30. *Nested-intervals. If a sequence of closed intervals $[a_n, b_n]$ is such that each member $[a_{n+1}, b_{n+1}]$ is contained in the preceding one $[a_n, b_n]$ and $\lim (b_n - a_n) = 0$, then there is one and only one point common to all the intervals of the sequence.*

Since each interval member of the sequence is contained in the preceding one, therefore, we have

$$a_1 \le a_2 \le \ldots \le a_n \le \ldots$$

$$b_1 \ge b_2 \ge \ldots \ge b_n \ge \ldots,$$

so that sequence $\{a_n\}$ is monotonic increasing, and $\{b_n\}$ is monotonic decreasing. Also $a_n \le b_1$ and $b_n \ge a_1$, for all n, so that sequence $\{a_n\}$ is bounded above by b_1 and $\{b_n\}$ bounded below by a_1.

Thus both the sequences are convergent.

Let $\lim a_n = \xi$ and $\lim b_n = \eta$.

Again

$$0 = \lim (b_n - a_n) = \lim b_n - \lim a_n = \xi - \eta$$

$$\Rightarrow \qquad \xi = \eta.$$

Obviously ξ is the upper bound of the sequence $\{a_n\}$ and lower bound of the sequence $\{b_n\}$, and hence

$$a_n \le \xi \le b_n, \quad \forall\, n$$

so that ξ belongs to all the intervals.

Let, if possible, ξ_1, ξ_2 be two different points common to all the intervals, and let $\xi_1 < \xi_2$.

Then

$$a_n \le \xi_1 < \xi_2 \le b_n, \quad \forall\, n$$

i.e.,

$$b_n - a_n \ge \xi_2 - \xi_1 \ne 0, \text{ for all } n,$$

which is a contradiction to the fact that $\xi_1 < \xi_2$ $\quad (\because \lim (b_n - a_n) = 0)$

Hence the result.

The following result is an important generalization of the theorem on nested intervals and is due to G. Cantor.

Theorem 31. *Cantor's intersection theorem for real line. If $F = \{F_n\}$ is a countable class of non-empty closed and bounded sets such that*

$$F_1 \supset F_2 \supset F_3 \ldots \supset F_n, \text{ then } \bigcap_{n=1}^{\infty} F_n \text{ is non-empty.}$$

Since each F_n is a non-empty closed and bounded set, therefore, there exist sequences of real numbers M_n and m_n belonging to F_n, such that

$$M_n = \sup F_n,\ m_n = \inf F_n,$$

then $M_n \ge M_{n+1}$ and $m_n \le m_{n+1}$, for each $n \in \mathbf{N}$. Now the lower bound for the set $\bigcap_{n=1}^{\infty} F_n$ is the lower bound for the sequence $\{M_n\}$ of upper bounds. Thus $\{M_n\}$ is a non-increasing sequence which is bounded below and therefore convergent.

Let $\lim_{n\to\infty} M_n = M$

we shall show that $M \in \bigcap_{n=1}^{\infty} F_n$. Let, if possible, $M \notin \bigcap_{n=1}^{\infty} F_n$. Then there will be at least one neighbourhood, say, $]M - \varepsilon, M + \varepsilon[$, $\varepsilon > 0$ which contains no point of $\bigcap_{n=1}^{\infty} F_n$

$\Rightarrow$ $]M - \varepsilon, M + \varepsilon[$ contains no point of F_n for some value of n, say, m.

$\Rightarrow$ $]M - \varepsilon, M + \varepsilon[$ contains no point of F_n, for $n \geq m$

$\Rightarrow$ $M_n \notin\,]M - \varepsilon, M + \varepsilon[, \ \forall\, n \geq m,$

contradicting the fact that $\{M_n\}$ converges to M.

Hence, $M \in \bigcap_{n=1}^{\infty} F_n$.

EXERCISE

1. Show that the sequence $\{S_n\}$, where

$$S_n = 1 + \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + \ldots + \frac{1}{(n-1)!}$$

is convergent.

2. Prove that the sequence $\{S_n\}$ defined by the recursion formula

$$S_{n+1} = \sqrt{7 + S_n}\,,\ S_1 = \sqrt{7}\,,$$

converges to the positive root of

$$x^2 - x - 7 = 0.$$

3. If $\{S_a\}$ be a sequence such that

$$S_{n+1} = 2 - \frac{1}{S_n},\ n \geq 1 \text{ and } S_1 = \frac{3}{2},$$

then show that the sequence $\{S_n\}$ is bounded and monotonic and converges to 1.

4. Given that $\{a_n\}$ is a sequence such that

$$a_2 \leq a_4 \leq a_6 \leq \ldots \leq a_5 \leq a_3 \leq a_1$$

and a sequence $\{b_n\}$, where $b_n = a_{2n-1} - a_{2n}$, converges to 0, then show that the sequence $\{a_n\}$ is convergent.

5. Let $\{a_n\}$ be a sequence, defined by

$$a_{n+1} = \frac{4 + 3a_n}{3 + 2a_n},\ n \geq 1\ \ a_1 = 1,$$

show that $\{a_n\}$ converges to $\sqrt{2}$.

6. If $\{a_n\}$ be a sequence of positive real numbers such that

$$a_n = \sqrt{a_{n-1}\,a_{n-2}}\,,\ n > 2,$$

then show that the sequence converges to $(a_1 a_2^{\,2})^{1/3}$.

CHAPTER 3

7. Let $\{a_n\}$ be a sequence defined by

$$a_{n+1} = \frac{1}{k}\left(a_n + \frac{k}{a_n}\right), k > 1 \text{ and } a_1 > 0,$$

show that $\{a_n\}$ converges to $\sqrt{\dfrac{k}{k-1}}$.

8. Show that the sequence $\{a_n\}$ defined by

$$a_{n+1} = 1 - \sqrt{1 - a_n}, \ \forall\, n \geq 1 \text{ and } a_1 < 1$$

converges to '0'.

9. Show that the sequences $\{a_n\}$ and $\{b_n\}$ defined by

$$a_{n+1} = \tfrac{1}{2}(a_n + b_n) \text{ and } b_{n+1} = \sqrt{a_n b_n},$$

converge to the common limit, where

$$a > b > 0 \text{ and } a_1 = \tfrac{1}{2}(a + b), b_1 = \sqrt{ab}.$$

(*Hint:* A.M. ≥ G.M., the sequences $\{a_n\}$ and $\{b_n\}$ are monotonically decreasing and increasing respectively).

10. If a_1 and b_1 are positive and if for all $n \geq 1$,

$$a_{n+1} = \frac{a_n + b_n}{2}, \frac{2}{b_{n+1}} = \frac{1}{a_n} + \frac{1}{b_n}$$

then show that $\{a_n\}$ and $\{b_n\}$ are monotonic sequences which converge to the common limit l, where $l^2 = a_1 b_1$.

11. If $\{a_{n_k}\}$ is a subsequence of the sequence $\{a_n\}$, then show that

$$\underline{\lim}\, a_n \leq \underline{\lim}\, a_{n_k} \leq \overline{\lim}\, a_{n_k} \leq \overline{\lim}\, a_n.$$

Deduce that, if the sequence $\{a_n\}$ converges to a, then all its subsequences will converge to the same limit a.

12. If $\{a_n\}$ is a bounded sequence, then show that there exists a subsequence $\{an_k\}$ of $\{a_n\}$, such that

$$\lim_{k\to\infty} a_{n_k} = \lim a_n$$

4 Infinite Series

1. INTRODUCTION

In this chapter we shall discuss the techniques of testing the *behaviour* of infinite series as regards convergence. The most important technique for series, all of whose terms are of the same sign (all positive or all negative), is to compare the given series with another suitably chosen series with known behaviour. So, first of all, comparison tests are discussed, and then some special tests for convergence are considered. *Leibnitz's test* for alternating series, *Abel's* and *Dirichlet's tests* for arbitrary term series, and *Dirichlet's* and *Riemann's* theorems on rearrangement of terms will be discussed in detail towards the end.

1.1 A *series* is the sum of the terms of a sequence. Thus if $u_1, u_2, u_3, \ldots$ is a sequence then the sum $u_1 + u_2 + u_3 + \ldots$ of all the terms is called an *infinite series* and is denoted by $\sum_{n=1}^{\infty} u_n$ or simply by $\sum u_n$.

Evidently we cannot just add up all the infinite number of terms of the series in the ordinary way and in fact it is not obvious that this kind of sum has any meaning. We thus start by associating with the given series, a sequence $\{S_n\}$, where S_n denotes the sum of the first n terms of the series. Thus,

$$S_n = u_1 + u_2 + \ldots + u_n, \quad \forall n.$$

The sequence $\{S_n\}$ is called the *sequence of partial sums* of the series and the partial sums, $S_1 = u_1$, $S_2 = u_1 + u_2$, $S_3 = u_1 + u_2 + u_3$, and so on, may be regarded as approximations to the full infinite sum $\sum_{n=1}^{\infty} u_n$ of the series. If the sequence $\{S_n\}$ of partial sums converges, then the series is regarded as convergent, and $\lim S_n$ is said to be the *sum of the series*. If, however, $\{S_n\}$ does not tend to a limit, we must take it that the sum of the infinite series does not exist. We express this fact by saying that the series does not converge. In fact *an infinite series is said to converge, diverge or oscillate according as its sequence of partial sums* $\{S_n\}$ *converges, diverges or oscillates*.

1.2 A Necessary Condition for Convergence

Theorem 1. *A necessary condition for convergence of an infinite series* $\sum u_n$ *is that* $\lim_{n\to\infty} u_n = 0$.

Let $S_n = u_1 + u_2 + \ldots + u_n$ so that $\{S_n\}$ is the sequence of partial sums.

Since the series converges, therefore, the sequence $\{S_n\}$ also converges.

Consequently

$$\lim_{n\to\infty} S_n = s(\text{say})$$

Now,

$$u_n = S_n - S_{n-1},\ \ n > 1$$

$$\therefore \qquad \lim_{n\to\infty} u_n = \lim_{n\to\infty} S_n - \lim_{n\to\infty} S_{n-1}$$

$$= s - s = 0$$

Hence for a convergent series,

$$\lim u_n = 0$$

In other words, a series cannot converge if its nth term does not tend to zero,

Notes:

1. It must be clearly understood that $\lim u_n = 0$ does not prove that a series in convergent, for there exist series which do not converge even though $\lim u_n = 0$. See Example 2.
2. However, $\lim u_n \neq 0$ proves that the series does not converge, see Example 1 below.

Example 1. Show that the series

$$\frac{1}{2} + \frac{2}{3} + \frac{3}{4} + \ldots$$

is not convergent.

■ Here

$$u_n = \frac{n}{n+1}$$

$$\therefore \qquad \lim u_n = \lim \frac{n}{n+1} = 1 \neq 0$$

Since $\lim u_n \neq 0$, therefore, the series is not convergent.

1.3 Cauchy's General Principle of Convergence for Series

A necessary and sufficient condition for the convergence of an infinite series $\sum\limits_{n=1}^{\infty} u_n$ is that the sequence of its partial sums $\{S_n\}$ is convergent.

Therefore, a test for convergence of infinite series may be derived from our knowledge of sequences.

Theorem 2. *A series $\sum u_n$ converges iff for each $\varepsilon > 0$, there exists a positive integer m such that*

$$\left| u_{n+1} + u_{n+2} + \ldots + u_{n+p} \right| < \varepsilon,\ \ \forall\ n \geq m \text{ and } p \geq 1$$

By Cauchy's General Principle of Convergence (for sequences), the sequence $\{S_n\}$of partial sums of $\sum u_n$ converges iff to each $\varepsilon > 0\ \exists$ a positive integer m, such that

$$\left| S_{n+p} - S_n \right| < \varepsilon,\ \ \forall\ n \geq m \text{ and } p \geq 1$$

or

$$\left|u_{n+1}+u_{n+2}+\ldots+u_{n+p}\right|<\varepsilon, \ \forall\, n\geq m \text{ and } p\geq 1$$

Example 2. Show that the series $\sum \frac{1}{n}$ does not converge.

- Suppose, if possible, the series converges.

  Therefore, for any given $\varepsilon\left(\text{say}, \frac{1}{4}\right)$, $\exists$ a positive integer m such that

$$\left|\frac{1}{n+1}+\frac{1}{n+2}+\ldots+\frac{1}{n+p}\right|<\varepsilon, \ \forall\, n\geq m \text{ and } p\geq 1$$

  In particular, if $n=m$ and $p=m$, we get

$$\frac{1}{n+1}+\frac{1}{n+2}+\ldots+\frac{1}{n+p}=\frac{1}{m+1}+\frac{1}{m+2}+\ldots+\frac{1}{2m}$$

$$>m\cdot\frac{1}{2m}=\frac{1}{2}>\varepsilon$$

  Thus, there is a contradiction. Hence, the given series does not converge.

  It may also be seen that $\lim u=\lim\frac{1}{n}=0$ even though the series does not converge.

Example 3. If $u_n>0$ and $\sum u_n$ is convergent, with the sum S, then prove that

$$\frac{u_n}{u_1+u_2+\ldots+u_n}<\frac{2u_n}{S},$$

when n is sufficiently large. Also prove that $\sum\frac{u_n}{u_1+u_2+\ldots+u_n}$ is convergent.

- Since $\sum u_n$ is convergent with the sum S, therefore, for $\varepsilon>0\ \exists$ a positive integer m, such that

$$|S_n-S|<\varepsilon \ \ \forall\, n\geq m, \text{ where } S_n=u_1+u_2+\ldots+u_n,$$

or

$$S-\varepsilon<S_n<S+\varepsilon, \ \forall\, n\geq m$$

  In particular for $\varepsilon=\frac{1}{2}S>0$,

$$\frac{1}{2}S<S_n<\frac{3}{2}S, \ \ \forall n\geq m$$

$$\Rightarrow \qquad \frac{2}{S}>\frac{1}{S_n}>\frac{2}{3S}, \ \forall\, n\geq m$$

$$\therefore \qquad \frac{u_n}{S_n}<\frac{2u_n}{S}, \ \forall\, n\geq m.$$

  Now, $\frac{u_{n+1}}{S_{n+1}}+\frac{u_{n+2}}{S_{n+2}}+\ldots+\frac{u_{n+p}}{S_{n+p}}<\frac{2}{S}\left(S_{n+p}-S_n\right), \ \forall\, n\geq m$ and $p\geq 1$.

CHAPTER 4

Since $\sum u_n$ is convergent, then for given $\varepsilon > 0$, there exists a positive integer m_1 such that

$$S_{n+p} - S_n < \frac{\varepsilon S}{2}, \ \forall \ n \geq m_1$$

$$\frac{u_{n+1}}{S_{n+1}} + \frac{u_{n+2}}{S_{n+2}} + \ldots + \frac{u_{n+p}}{S_{n+p}} < \frac{2}{S} \cdot \frac{\varepsilon S}{2} = \varepsilon, \quad \forall \ n \geq \max\left(m, m_1\right)$$

Hence by *Cauchy's General Principle of Convergence*, $\sum \dfrac{u_n}{u_1 + u_2 + \ldots + u_n}$ is convergent.

1.4 Some Preliminary Theorems

Theorem 3. *If $\Sigma u_n = u$, then $\Sigma c u_n = cu$, independent of n.*

The result follows at once from the identity

$$\sum_{r=1}^{n} c u_r = c \sum_{r=1}^{n} u_r$$

on making n tend to infinity.

Theorem 4. If $\sum\limits_{n=1}^{\infty} u_n = u$, *then* $\sum\limits_{n=0}^{\infty} u_n = u + u_0$, *and* $\sum\limits_{n=2}^{\infty} u_n = u - u_1$.

Let $S'_n = \sum\limits_{r=0}^{n} u_r$ and $S_n = \sum\limits_{r=1}^{n} u_r$.

Clearly

$$S'_n = u_0 + S_n$$

$\therefore$ By letting n tend to infinity

$$\sum_{n=0}^{\infty} u_n = u_0 + u$$

The proof of the second part is similar.

A slight modification and extension enables us to conclude that the insertion or removal of any finite number of terms from a convergent series does not affect its convergence. Of course, the sums of the various series are related in the expected way.

It is also clear that if the series $\sum\limits_{n=1}^{\infty} u_n$ is divergent, the changed series $\sum\limits_{n=1}^{\infty} c u_n$, $\sum\limits_{n=0}^{\infty} u_n$ or $\sum\limits_{n=2}^{\infty} u_n$ is also divergent.

Hence, the behaviour of a series as regards convergence is not altered by

(*i*) *the alteration, addition or omission of a finite number of terms; or*

(*ii*) *multiplication of all the terms by a finite number other than zero.*

Theorem 5. ***Convergent series may be added or subtracted term by term.*** *If $\sum u_n = u$ and $\sum v_n = v$ then $\sum w_n = u \pm v$, where $w_n = u_n \pm v_n$, for all n.*

The result follows from the identity

$$\sum_{r=1}^{n} w_r = \sum_{r=1}^{n} (u_r \pm v_r) = \sum_{r=1}^{n} u_r \pm \sum_{r=1}^{n} v_r$$

by making n tend to infinity.

The same proof shows that

(*i*) if any two of the three series are convergent, the third is also convergent,

(*ii*) if one of the series is divergent and another convergent then the third is necessarily divergent, but

(*iii*) if two of the series are divergent, no conclusion can be drawn about the behaviour of the third, which may converge or diverge.

Theorem 6. *If a series $\sum u_n$ converges to the sum u then so does any series obtained from $\sum u_n$ by grouping the terms in brackets without altering the order of the terms.*

Suppose that the series derived from $\sum u_n$ by the insertion of brackets is $\sum v_n$ and let σ_r denote the rth partial sum of the series $\sum v_n$. Suppose that σ_r contains n_r terms of the given series then since the order of the terms is unaltered, $\sigma_r = S_{n_r}$.

Also as $r \to \infty$, $n_r \to \infty$.

Since the given series converges to u, the sequence $\{S_n\}$ of its partial sums also converges to u.

Hence, as $r \to \infty$ or $n_r \to \infty$, $s_{n_r} \to u$, and hence $\sigma_r \to u$.

Remarks:

1. Converse of the theorem is not always true.

   For example, the series $(1 - 1) + (1 - 1) + \ldots$ is convergent, whereas the series $1 - 1 + 1 - 1 + \ldots$, obtained by removing the brackets is not.

   Hence, in convergent series brackets may be *inserted* at will without affecting convergence but may not be removed.

   In the case of convergent positive term series, or the absolutely convergent series, however, it will be shown later that the brackets may be inserted or removed without affecting convergence.

2. The theorem may as well be proved by the following alternative argument.

   Since the given series converges to u, its sequence $\{S_n\}$ of partial sums will also converge to u. Therefore, the sequence $\{\sigma_n\}$ of partial sums of $\sum v_n$, being a subsequence of $\{S_n\}$ will also converge to u.

3. The theorem may be restated, "*A series obtained from a given convergent series by a grouping of terms converges to the same limit*".

   By *grouping* we simply mean the placing of brackets or associating the terms of the series without changing the order of the terms.

2. POSITIVE TERM SERIES

Series with non-negative terms are the simplest and the most important type of series one comes across. The simplicity arises mainly from the fact that the sequence of its partial sums is monotonic increasing.

Let $\sum u_n$ be an infinite series of positive terms and $\{S_n\}$, the sequence of its partial sums, so that

$$S_n = u_1 + u_2 + \ldots + u_n \geq 0, \ \forall\, n$$

$$\therefore \qquad S_n - S_{n-1} = u_n \geq 0$$

$$\Rightarrow \qquad S_n \geq S_{n-1}, \quad \forall\, n > 1$$

Thus the sequence $\{S_n\}$ of partial sums of a series of positive terms is a monotonic increasing sequence.

Since a monotonic increasing sequence can either converge, or diverge to ∞, but cannot oscillate, therefore, there are only two possibilities for a positive term series—it may either converge or diverge to $+\infty$.

Theorem 7. *A positive term series converges iff the sequence of its partial sums is bounded above.*

We know that the sequence of partial sums of a positive term series is a monotonic increasing sequence and a monotonic increasing sequence converges iff it is bounded above. Therefore, it follows that a positive term series converges iff its sequence of partial sums is bounded above.

Remarks:

1. The sequence of partial sums of a series with negative terms can be shown to be monotonic decreasing and hence a series with negative terms converges iff the sequence of its partial sums is bounded below.
2. It may similarly be seen that a series of negative terms can either converge, or diverge to $-\infty$.
3. A series of positive terms can either converge, or diverge to $+\infty$. But a series with arbitrary terms can have five possible behaviours depending upon the behaviour of the sequence of its partial sums.

2.1 A Necessary Condition for Convergence of Positive Term Series

We know that for any convergent series, the nth term $u_n \to 0$ as $n \to \infty$, we now give another necessary condition which holds for positive term series only.

Theorem 8. ***Pringsheim's Theorem.*** *If a series Σu_n of positive monotonic decreasing terms converges then not only $u_n \to 0$ but also $nu_n \to 0$ as $n \to \infty$.*

We know that for a convergent series, for any $\varepsilon > 0$, a positive integer N exists such that

$$\left| u_{m+1} + u_{m+2} + \ldots + u_{m+p} \right| < \frac{\varepsilon}{2}, \quad \forall\, m \geq N,\ p \geq 1$$

Let us choose

$$m + p = n > 2N$$

and

$$m = \left[\frac{n}{2}\right]$$

the greatest integer not greater than $\dfrac{n}{2}$

$$\therefore \qquad u_{m+1} + u_{m+2} + \ldots + u_n < \frac{\varepsilon}{2}$$

But $\Sigma\, u_n$ is positive monotonic decreasing,

$$\therefore \qquad (n - m)\, u_n < u_{m+1} + u_{m+2} + \ldots + u_n < \frac{\varepsilon}{2}$$

or

$$\frac{1}{2}nu_n < \frac{1}{2}\varepsilon$$

i.e.,

$$nu_n < \varepsilon, \ \forall n \geq N$$

$\therefore$

$$\lim nu_n = 0$$

Notes:

1. The condition $u_n \to 0$ holds for all types of convergent series but a convergent series of positive monotonic decreasing terms satisfies the additional condition, $nu_n \to 0$.
2. The condition $nu_n \to 0$ is only a necessary not a sufficient condition for the convergence of the present type of series. If nu_n does not tend to zero then the series $\sum u_n$ is certainly divergent, *e.g.*, the harmonic series $\sum \frac{1}{n}$ must diverge because it has positive monotonic decreasing terms and $n \cdot \frac{1}{n}$ does not tend to zero. However, $nu_n \to 0$ does not imply anything as to the possible convergence of $\sum u_n$, e.g., Abel's series $\sum \frac{1}{n \log n}$ diverges although it has positive monotonic decreasing terms, and $nu_n \to 0$.

2.2 Geometric Series

The positive term geometric series $1 + r + r^2 + \ldots$, converges for $r < 1$, and diverges to $+\infty$ for $r \geq 1$.

Case I. $0 \leq r < 1$.

Let $\{S_n\}$ be the sequence of its partial sums, so that

$$S_n = 1 + r + r^2 + \ldots + r^{n-1}$$

i.e.,

$$S_n = \frac{1 - r^n}{1 - r} = \frac{1}{1 - r} - \frac{r^n}{1 - r} \leq \frac{1}{1 - r}, \ \forall n,$$

so that $\{S_n\}$ is bounded above and hence the series converges.

Case II. When $r = 1$, $S_n = n$, so that $\{S_n\}$ is not bounded above and hence the series diverges to $+\infty$.

Case III. When $r > 1$, every term of S_n after the first is greater than 1, so that

$$S_n > n, \ \forall n$$

Therefore, the sequence $\{S_n\}$ is not bounded above and consequently the given series diverges to $+\infty$.

Hence, the given series converges if $r < 1$ and diverges if $r \geq 1$.

2.3 A Comparison Series

As mentioned earlier, an important technique for testing the convergence of a series is to compare the given series with a suitably selected series with known behaviour. We now discuss one such series which is most frequently used for such a purpose.

Theorem 9. *A positive term series $\Sigma \frac{1}{n^p}$ is convergent if and only if $p > 1$.*

Let S_n denote the sum of the first n terms of the given series.

Case I. When $p > 1$.

Now

$$\frac{1}{1^p} = 1$$

$$\frac{1}{2^p} + \frac{1}{3^p} < \frac{2}{2^p} = \frac{1}{2^{p-1}}$$

$$\frac{1}{4^p} + \frac{1}{5^p} + \frac{1}{6^p} + \frac{1}{7^p} < \frac{4}{4^p} = \frac{1}{4^{p-1}} = \left(\frac{1}{2^{p-1}}\right)^2$$

$$\frac{1}{8^p} + \frac{1}{9^p} + \dots + \frac{1}{15^p} < \frac{8}{8^p} = \frac{1}{8^{p-1}} = \left(\frac{1}{2^{p-1}}\right)^3$$

$$\dots \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots$$

$$\dots \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots$$

$$\frac{1}{(2^n)^p} + \frac{1}{(2^n+1)^p} + \dots + \frac{1}{(2^{n+1}-1)^p} < \frac{2^{n+1}-2^n}{(2^n)^p} = \frac{1}{(2^n)^{p-1}} = \left(\frac{1}{2^{p-1}}\right)^n$$

Adding,

$$S_{2^{n+1}-1} < 1 + \frac{1}{2^{p-1}} + \left(\frac{1}{2^{p-1}}\right)^2 + \dots + \left(\frac{1}{2^{p-1}}\right)^n$$

$$= \frac{1 - \left(\frac{1}{2^{p-1}}\right)^{n+1}}{1 - \frac{1}{2^{p-1}}} = \frac{2^{p-1}\left[1 - \left(\frac{1}{2^{p-1}}\right)^{n+1}\right]}{2^{p-1} - 1}$$

$$< \frac{2^{p-1}}{2^{p-1}-1}, \quad \forall\, n$$

We know that when n is any positive integer,

$$2^{n+1} - 1 > 2^n > n$$

$$\therefore \qquad S_n < S_{2^n} < S_{2^{n+1}-1} < \frac{2^{p-1}}{2^{p-1}-1}$$

Since for a given p, $\frac{2^{p-1}}{2^{p-1}-1}$ is a fixed number, therefore, the sequence $\{S_n\}$ of partial sums of the given positive term series is bounded above and hence the series converges for $p > 1$.

Case II. When $p \le 1$.

We know, if n is any positive integer and $p \le 1$, then

$$n^p \le n \Rightarrow \frac{1}{n^p} \ge \frac{1}{n}$$

$\therefore$
$$1 + \frac{1}{2^p} \ge 1 + \frac{1}{2} > \frac{1}{2}$$

$$\frac{1}{3^p} + \frac{1}{4^p} \ge \frac{1}{3} + \frac{1}{4} > \frac{2}{4} = \frac{1}{2}$$

$$\frac{1}{5^p} + \frac{1}{6^p} + \frac{1}{7^p} + \frac{1}{8^p} \ge \frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8} > \frac{4}{8} = \frac{1}{2}$$

$$\frac{1}{9^p} + \frac{1}{10^p} + \ldots + \frac{1}{16^p} \ge \frac{1}{9} + \frac{1}{10} + \ldots + \frac{1}{16} > \frac{8}{16} = \frac{1}{2}$$

$$\ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots$$

$$\ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots$$

$$\frac{1}{\left(2^{m-1}+1\right)^p} + \frac{1}{\left(2^{m-1}+2\right)^p} + \ldots + \frac{2}{\left(2^m\right)^p} \ge \frac{1}{2^{m-1}+1} + \frac{1}{2^{m-1}+2} + \ldots + \frac{1}{2^m} > \frac{2^m - 2^{m-1}}{2^m} = \frac{1}{2}$$

Adding,

$$S_{2^m} > \frac{m}{2}$$

We shall now show that $\{S_n\}$ is not bounded above.

If G be any number, however, large, then $\exists\, m \in \mathbf{N}$ such that

$$\frac{m}{2} > G$$

Let $n > 2^m$,

$\therefore$
$$S_n > S_{2^m} > G$$

Thus, the sequence $\{S_n\}$ of partial sums of the given series is not bounded above, and hence the series diverges for $p \le 1$.

Thus, the given series $\Sigma \frac{1}{n^p}$ converges iff $p > 1$.

3. COMPARISON TESTS FOR POSITIVE TERM SERIES

Two types of comparison tests shall now be discussed. In the first type, the general term of one series will be compared with the general term of the second series. In the second type, the ratio of two consecutive terms of one series will be compared to the ratio of the corresponding consecutive terms of the second series.

3.1 Comparison Test (*First type*)

I. *If Σu_n and Σv_n are two positive term series, and $k \neq 0$, a fixed positive real number (independent of n) and there exists a positive integer m such that $u_n \leq kv_n, \;\forall\, n \geq m$, then*

(*i*) *Σu_n is convergent, if Σv_n is convergent, and*

(*ii*) *Σv_n is divergent, if Σu_n is divergent.*

Let $n \geq m$, $S_n = u_1 + u_2 + ... + u_n$, and $t_n = v_1 + v_2 + ... + v_n$.

Now for all $n \geq m$, we have

$$S_n - S_m = u_{m+1} + u_{m+2} + ... + u_n$$
$$\leq k\left(v_{m+1} + v_{m+2} + ... + v_n\right) = k(t_n - t_m)$$

or

$$S_n \leq kt_n + (S_m - kt_m)$$

$$\Rightarrow \qquad S_n \leq kt_n + h \qquad ...(1)$$

where $h = S_m - kt_m$, is a finite quantity.

(*i*) If Σv_n is convergent, then the sequence $\{t_n\}$ of its partial sums is bounded above, so that $\exists$ a number B such that

$$t_n \leq B, \;\forall\, n$$

So from (1), we get

$S_n \leq kB + h$, for all $n \geq m$,

$\Rightarrow$ the sequence $\{S_n\}$ is bounded above.

$\Rightarrow$ Σu_n is convergent.

(*ii*) If Σu_n is divergent, then the sequence $\{S_n\}$ of its partial sums is not bounded above, so that if G be any number, however, large, $\exists$ a positive integer m_0 such that

$$S_n > G \;\;\forall\, n \geq m_0.$$

Thus from (1), $\forall\, n \geq \max\,(m, m_0)$,

$$t_n \geq \frac{1}{k}(G - h),\; k \neq 0$$

$\Rightarrow$ the sequence $\{t_n\}$ is unbounded

$\Rightarrow$ Σv_n is divergent.

II. Limit Form. *If Σu_n* and *Σv_n are two positive term series such that $\lim\limits_{n\to\infty}\left(\dfrac{u_n}{v_n}\right) = l$, where l is a non-zero finite number, then the two series converge or diverge together.*

Evidently $l > 0$.

Let ε be a positive number such that $l - \varepsilon > 0$.

Since $\lim\limits_{n \to \infty} \dfrac{u_n}{v_n} = l$, therefore $\exists$ a positive integer m such that

$$\left| \frac{u_n}{v_n} - l \right| < \varepsilon, \ \forall \, n \geq m$$

$$\Rightarrow \qquad l - \varepsilon < \frac{u_n}{v_n} < 1 + \varepsilon, \ \forall \, n \geq m$$

$$\Rightarrow \qquad (l - \varepsilon) v_n < u_n < (1 + \varepsilon) v_n, \ \forall \, n \geq m \qquad ...(1)$$

Now, if Σv_n is convergent, then from (1)

$$u_n < (1 + \varepsilon) v_n, \ \forall \, n \geq m$$

so that by Test I, Σu_n is convergent.

Again, if Σv_n is divergent, then from (1)

$$u_n > (1 - \varepsilon) v_n, \ \forall \, n \geq m$$

so that by Test I, Σu_n is divergent.

Similarly, we may show that Σv_n converges or diverges with Σu_n. Hence, the two series behave alike.

3.2 Comparison Test (*Second type*)

III. *If Σu_n and Σv_n are two positive term series and $\exists$ a positive integer m such that*

$$\frac{u_n}{u_{n+1}} \geq \frac{v_n}{v_{n+1}}, \ \forall \, n \geq m,$$

then (i) Σu_n is convergent, if Σv_n is convergent, and (ii) Σv_n is divergent, if Σu_n is divergent.

Let $S_n = u_1 + u_2 + ... + u_n$

and $\quad t_n = v_1 + v_2 + ... + v_n$

For $n \geq m$, we have

$$\frac{u_m}{u_n} = \frac{u_m}{u_{m+1}} \cdot \frac{u_{m+1}}{u_{m+2}} \cdots \frac{u_{n-1}}{u_n} \geq \frac{v_m}{v_{m+1}} \cdot \frac{v_{m+1}}{v_{m+2}} \cdots \frac{v_{n-1}}{v_n} = \frac{v_m}{v_n}$$

$$\Rightarrow \qquad u_n \leq \frac{u_m}{v_m} v_n$$

Since m is a fixed positive integer, therefore u_m/v_m is fixed number, say k. Thus $\forall \, n \geq m$ we have

$$u_n \leq k v_n$$

Hence by Test I, Σu_n converges if Σv_n converges and Σv_n diverges if Σu_n diverges.

CHAPTER 4

Notes:

1. For practical purposes, Test II is very useful and can be easily applied.
2. For a successful application of the comparison test, we first make an estimate of the magnitude of the general term u_n of the given series, and then select the auxiliary series Σv_n of such a magnitude that $\lim\ (u_n/v_n) = l \neq 0, \infty$, or in other words $u_n \sim v_n$. Thus for large values of n,

$$\sqrt{n^3+1} \sim n^{3/2}, \frac{n^r}{(1+n)^s} \sim n^{r-s}$$

$$\sin\frac{1}{n} \sim \frac{1}{n}$$

3. It will help to remember that for large n, $e^{an} \gg n^b \gg (log\ n)^c$, where a, b, c are positive numbers.

3.3 Solved Examples

Example 4. Show that the series $1 + \frac{1}{2!} + \frac{1}{3!} + \frac{1}{4!} + \ldots$ is convergent.

■ We have

$$\frac{1}{2!} = \frac{1}{2}$$

$$\frac{1}{3!} < \frac{1}{2^2}$$

$$\frac{1}{4!} < \frac{1}{2^3}$$

$$\ldots \quad \ldots$$

$$\ldots \quad \ldots$$

$$\frac{1}{n!} < \frac{1}{2^{n-1}}$$

$$\therefore \quad 1 + \frac{1}{2!} + \frac{1}{3!} + \ldots + \frac{1}{n!} + \ldots < 1 + \frac{1}{2} + \frac{1}{2^2} + \ldots$$

Thus each term of the given series after the second is less than the corresponding term of the convergent geometric series

$$1 + \frac{1}{2} + \frac{1}{2^2} + \ldots$$

Thus by Test I, the given series converges.

Example 5. Show that the series

$$\frac{1}{(\log 2)^p} + \frac{1}{(\log 3)^p} + \ldots + \frac{1}{(\log n)^p} + \ldots$$

diverges for $p > 0$.

■ Since $\lim_{n\to\infty} \frac{(\log n)^p}{n} = 0$,

$$\Rightarrow \qquad (\log n)^p < n, \ \forall\, n > 1$$

$$\therefore \qquad \frac{1}{(\log n)^p} < \frac{1}{n}, \ \forall\, n > 1$$

Let us compare the given series with the divergent series

$$\frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \ldots$$

Since each term of the given series is greater than the corresponding term of the divergent series, therefore, the given series diverges.

Example 6. Show that the series

$$\frac{1 \cdot 2}{3^2 \cdot 4^2} + \frac{3 \cdot 4}{5^2 \cdot 6^2} + \frac{5 \cdot 6}{7^2 \cdot 8^2} + \ldots \text{ is convergent}$$

■ Let us denote the given series by Σu_n, where

$$u_n = \frac{(2n-1)(2n)}{(2n+1)^2 (2n+2)^2}, \left(\sim \frac{1}{n^2}\right).$$

Let us compare it with the convergent series Σv_n, where $v_n = 1/n^2$.

Now

$$\lim_{n\to\infty} \frac{u_n}{v_n} = \lim_{n\to\infty} \frac{(2-1/n)(2)}{(2+1/n)^2 (2+2/n)^2} = \frac{1}{4}$$

Thus, the two series converge or diverge together.

Since Σv_n converges, therefore Σu_n also converges.

Example 7. Investigate the behaviour of the series whose nth term is $\sin 1/n$.

■ Let $u_n = \sin 1/n$ and $v_n = 1/n$.

Now

$$\lim_{n\to\infty} \frac{u_n}{v_n} = \lim_{n\to\infty} \frac{\sin 1/n}{1/n} = 1$$

Therefore, the two series behave alike.

Since Σv_n diverges, therefore $\Sigma \sin 1/n$ also diverges.

Example 8. Test for convergence of the series whose nth term is

$$\{(n^3+1)^{1/3} - n\}$$

■ Let

$$u_n = \left(n^3+1\right)^{1/3} - n$$

$$= n\left\{\left(1+\frac{1}{n^3}\right)^{1/3} - 1\right\}$$

CHAPTER 4

$$= n\left\{\frac{1}{3n^3} + \ldots\right\} = \frac{1}{3n^2} + \ldots \left(\sim \frac{1}{n^2}\right)$$

and

$$\Sigma v_n = \Sigma \frac{1}{n^2}$$

Now

$$\lim_{n\to\infty} \frac{u_n}{v_n} = \frac{1}{3}$$

Therefore, the two series converge and diverge together.

Since Σv_n converges, therefore, the given series also converges.

Example 9. Test the convergence of the series $\Sigma \dfrac{1}{n^{1+1/n}}$.

■ Let

$$u_n = \Sigma \frac{1}{n^{1+1/n}} \text{ and } v_n = \frac{1}{n}$$

Now

$\therefore$

$$\lim_{n\to\infty} \frac{u_n}{v_n} = \lim_{n\to\infty} \frac{1}{n^{1/n}} = 1$$

Hence, the two series Σu_n and Σv_n behave alike.

Since Σv_n is divergent, therefore $\Sigma \dfrac{1}{n^{1+1/n}}$ is also divergent.

EXERCISE

Test the convergence of the following series:

1. $\dfrac{1}{1\cdot 2\cdot 3} + \dfrac{3}{2\cdot 3\cdot 4} + \dfrac{5}{3\cdot 4\cdot 5} + \dfrac{7}{4\cdot 5\cdot 6} + \ldots$
2. $\dfrac{1}{\sqrt{1\cdot 2}} + \dfrac{1}{\sqrt{2\cdot 3}} + \dfrac{1}{\sqrt{3\cdot 4}} + \ldots$
3. $\dfrac{1}{4\cdot 6} + \dfrac{\sqrt{3}}{6\cdot 8} + \dfrac{\sqrt{5}}{8\cdot 10} + \dfrac{\sqrt{7}}{10\cdot 12} + \ldots$
4. $\Sigma \dfrac{n+1}{n^p}$
5. $\Sigma \dfrac{1}{\sqrt{n} + \sqrt{n+1}}$
6. $\Sigma\left(\sqrt{n^4+1} - \sqrt{n^4-1}\right)$ [*Hint*: Rationalize]
7. (*i*) $\Sigma \sin \dfrac{1}{n^2}$, (*ii*) $\Sigma \cos \dfrac{1}{n}$

8. $\Sigma \dfrac{1}{\sqrt{n}} \tan \dfrac{1}{n}$

9. $\sum\limits_{n=1}^{\infty} e^{-n^2}$

10. $\sum\limits_{n=1}^{\infty} \dfrac{5^n + 5}{3^n + 2}$

11. $\sum\limits_{n=2}^{\infty} \dfrac{1}{n^2 \log n}$

12. $\Sigma \left\{ \dfrac{\sqrt{n+1} - \sqrt{n-1}}{n} \right\}$

13. Show that the series $1 + \dfrac{1}{2^2} + \dfrac{1}{3^3} + \dfrac{1}{4^4} + \ldots$ converges.

[*Hint*: $n^n > 2^n$, for $n > 2$. Compare with the convergent geometric series $\Sigma 1/2^n$]

ANSWERS

1. Convergent, 2. Divergent, 3. Convergent, 4. Convergent, for, $p > 2$, 5. Divergent, 6. Convergent, 7. (*i*) Convergent, (*ii*) Divergent, 8. Convergent, 9. Convergent, 10. Divergent, 11. Convergent, 12. Convergent.

4. CAUCHY'S ROOT TEST

If Σu_n is a positive term series, such that $\lim\limits_{n\to\infty} (u_n)^{1/n} = l$, then the series

(*i*) *converges, if $l < 1$,*
(*ii*) *diverges, if $l > 1$, and*
(*iii*) *the test fails to give any definite information, if $l = 1$.*

Case I. $l < 1$.

Let us select a positive number ε, such that $l + \varepsilon < 1$.

Let $l + \varepsilon = \alpha < 1$.

Since, $\lim\limits_{n\to\infty} (u_n)^{1/n} = l$, therefore $\exists$ a positive integer m such that

$$\left| (u_n)^{1/n} - l \right| < \varepsilon, \ \forall \ n \geq m$$

$$\Rightarrow \quad l - \varepsilon < (u_n)^{1/n} < l + \varepsilon, \ \forall \ n \geq m$$

$$\Rightarrow \quad (l - \varepsilon)^n < u_n < (l + \varepsilon)^n = \alpha^n, \ \forall \ n \geq m$$

$$\Rightarrow \quad u_n < \alpha^n, \ \forall \ n \geq m.$$

CHAPTER 4

But since $\sum \alpha^n$ is a convergent geometric series (common ratio $\alpha < 1$), therefore, by comparison test, the series $\sum u_n$ converges.

Case II. $l > 1$.

Let us select a positive number ε such that $l - \varepsilon > 1$

Let $l - \varepsilon = \beta > 1$.

Since $\lim_{n \to \infty} (u_n)^{1/n} = l$, therefore $\exists$ a positive integer m_1 such that

$$l - \varepsilon < (u_n)^{1/n} < l + \varepsilon, \quad \forall\ n \geq m_1$$

$$\Rightarrow \qquad (l - \varepsilon)^n < u_n < (l - \varepsilon)^n, \quad \forall\ n \geq m_1$$

$$\Rightarrow \qquad u_n > (l - \varepsilon)^n = \beta^n, \quad \forall\ n \geq m_1.$$

But since $\sum \beta^n$ is a divergent geometric series (common ratio $\beta > 1$), therefore by comparison test, the series $\sum u_n$ diverges.

Note: The test fails to give any definite information for $l = 1$.

Consider the two series $\sum (1/n)$ and $\Sigma (1/n^2)$.

$\sum (1/n)$ diverges when $\lim_{n\to\infty} (1/n)^{1/n} = 1$, and $\sum (1/n^2)$ converges when $\lim_{n\to\infty} (1/n^2)^{1/n} = 1$.

Example 10. Test for convergence of the series whose general term is $\left(1 + \dfrac{1}{\sqrt{n}}\right)^{-n^{3/2}}$

■ Let $u_n = \dfrac{1}{\left(1 + 1/\sqrt{n}\right)^{n^{3/2}}}$, then

$$\lim_{n\to\infty} (u_n)^{1/n} = \lim_{n\to\infty} \frac{1}{\left(1 + 1/\sqrt{n}\right)^{\sqrt{n}}} = \frac{1}{e} < 1.$$

Hence, the series converges.

5. D'ALEMBERT'S RATIO TEST

If $\sum u_n$ is a positive term series, such that $\lim_{n\to\infty} \dfrac{u_{n+1}}{u_n} = l$, then the series

(*i*) *converges, if $l < 1$,*

(*ii*) *diverges, if $l > 1$, and*

(*iii*) *the test fails, if $l = 1$.*

Case I. $0 < l < 1$.

Let us select a positive number ε, such that $l + \varepsilon < 1$.

Let $l + \varepsilon = \alpha < 1, \alpha \neq 0$.

Since $\lim\limits_{n\to\infty} \dfrac{u_{n+1}}{u_n} = l$, therefore $\exists$ a positive integer m such that

$$\left|\frac{u_{n+1}}{u_n} - l\right| < \varepsilon, \ \forall\ n \geq m$$

$$\Rightarrow \qquad l - \varepsilon < \frac{u_{n+1}}{u_n} < l + \varepsilon, \ \forall\ n \geq m$$

$$\Rightarrow \qquad \frac{u_{n+1}}{u_n} < l + \varepsilon = \alpha, \ \forall\ n \geq m$$

For $n \geq m$,

$$\frac{u_n}{u_m} = \frac{u_{m+1}}{u_m} \cdot \frac{u_{m+2}}{u_{m+1}} \cdots \frac{u_n}{u_{n-1}} < \alpha^{n-m}$$

$$\Rightarrow \qquad u_n < \left(\frac{u_m}{\alpha^m}\right)\alpha^n, \ \forall\ n \geq m, \alpha < 1.$$

Since m is a fixed integer, therefore $\left(\dfrac{u_m}{\alpha^m}\right)$ is a fixed finite number, say k.

Thus, $\forall\ n \geq m$, we have

$$u_n < k\alpha^n$$

But since $\sum \alpha^n$ is a convergent geometric series (common ratio, $\alpha < 1$), therefore by comparison test $\sum u_n$ converges.

Case II. $l > 1$.

Let us select a positive number ε, such that $l - \varepsilon > 1$.

Let $l - \varepsilon = \beta > 1$.

Since $\lim\limits_{n\to\infty} \dfrac{u_{n+1}}{u_n} = l$, therefore $\exists$ a positive integer m_1 such that

$$l - \varepsilon < \frac{u_{n+1}}{u_n} < l + \varepsilon, \ \forall\ n \geq m_1$$

$$\Rightarrow \qquad \frac{u_{n+1}}{u_n} > l - \varepsilon = \beta, \ \forall\ n \geq m_1$$

CHAPTER 4

Now, for $n \geq m_1$,

$$\frac{u_n}{u_{m_1}} = \frac{u_{m_1+1}}{u_{m_1}} \cdot \frac{u_{m_1+2}}{u_{m_1+1}} \cdots \frac{u_n}{u_{n-1}} > \beta^{n-m_1}$$

$\Rightarrow$
$$u_n > \frac{u_{m_1}}{\beta^{m_1}} \beta^n, \ \forall \ n \geq m_1$$

Since m_1 is a fixed integer, therefore u_{m_1}/β^{m_1} is a fixed finite number, say k_1.

Thus, for $n \geq m_1$, we have

$$u_n > k_1 \beta^n$$

But $\Sigma \beta^n$ is a divergent geometric series (common ratio, $\beta > 1$), therefore by comparison test, Σu_n diverges.

Note: The test fails for $l = 1$ in the sense that it fails to give any definite information.

For example, consider the two series $\Sigma(1/n)$ and $\Sigma(1/n^2)$

$\Sigma \frac{1}{n}$ diverges when $\lim\limits_{n\to\infty} \frac{u_{n+1}}{u_n} = \lim\limits_{n\to\infty} \frac{n}{n+1} = 1$, and $\Sigma \frac{1}{n^2}$ converges when $\lim\limits_{n\to\infty} \frac{u_{n+1}}{u_n} = \lim\limits_{n\to\infty} \left(\frac{n}{n+1}\right)^2 = 1$.

Remark: Cauchy's Root Test is stronger than D' Alembert's Ratio Test and may succeed where Ratio-Test fails. For example, take the series Σu_n, where $u_{2n-1} = 1/2^{2n-1}$ and $u_{2n} = 1/3^{2n}$, $\forall \ n$.

Example 11. Test for convergence of the series $\Sigma \frac{n^2-1}{n^2+1} x^n$, $x > 0$.

■ Let $u_n = \frac{n^2-1}{n^2+1} x^n$.

$\therefore$
$$\lim_{n\to\infty} \frac{u_{n+1}}{u_n} = \lim_{n\to\infty} \frac{n^2+1}{n^2-1} \cdot \frac{(n+1)^2-1}{(n+1)^2+1} \cdot \frac{x^{n+1}}{x^n} = x$$

Hence by D'Alembert's Ratio Test the series converges if $x < 1$ and diverges if $x > 1$.

The test fails to give any information when $x = 1$.

When $x = 1$, $u_n = \frac{n^2-1}{n^2+1}$, and $\lim\limits_{n\to\infty} u_n = 1 \neq 0$

$\Rightarrow$ The series is divergent.

Hence, the series converges if $x < 1$ and diverges if $x \geq 1$.

6. RAABE'S TEST

If Σu_n is a positive term series, such that $\lim\limits_{n\to\infty} n\left(\frac{u_n}{u_{n+1}} - 1\right) = 1$, then the series

(i) *converges, if* $l > 1$,

(ii) *diverges, if* $l < 1$, and

(iii) *the test fails, if* $l = 1$.

Case I. $l > 1$.

Let us select a positive number ε, such that $l - \varepsilon > 1$.

Let $l - \varepsilon = \alpha > 1$.

Since $\lim_{n\to\infty} n\left(\frac{u_n}{u_{n+1}} - 1\right) = l$, therefore $\exists$ a positive integer m such that for all $n \geq m$,

$$l - \varepsilon < n\left(\frac{u_n}{u_{n+1}} - 1\right) < l + \varepsilon$$

$$\Rightarrow \qquad n\left(\frac{u_n}{u_{n+1}} - 1\right) > l - \varepsilon = \alpha$$

$$\Rightarrow \qquad nu_n - nu_{n+1} > \alpha u_{n+1}$$

$$\Rightarrow \qquad nu_n - (n+1)u_{n+1} > (\alpha - 1)u_{n+1}, \quad \alpha - 1 > 0$$

Putting $n = m, m + 1, m + 2, ..., n - 1$ and adding, we get

$$mu_m - nu_n > (\alpha - 1)\left(u_{m+1} + u_{m+2} + ... + u_n\right)$$

$$= (\alpha - 1)(S_n - S_m), \text{ where } S_n = \sum_{r=1}^{n} u_r$$

$$\Rightarrow \qquad (\alpha - 1)\left(S_n - S_m\right) < mu_m, \quad \forall\, n \geq m$$

$$\Rightarrow \qquad S_n < S_m + \frac{m}{\alpha - 1}u_m, \quad \forall\, n \geq m,$$

Since m is a fixed integer, therefore $S_m + \frac{m}{\alpha - 1}u_m$ is a fixed finite number.

Thus, the sequence $\{S_n\}$ of partial sums of the given series is bounded above and hence the series Σu_n is convergent.

Case II. $l < 1$.

Let us select a positive number ε, such that $l + \varepsilon < 1$.

Since $\lim_{n\to\infty} n\left(\frac{u_n}{u_{n+1}} - 1\right) = l$, therefore $\exists$ a positive integer m such that for all $n \geq m$,

$$l - \varepsilon < n\left(\frac{u_n}{u_{n+1}} - 1\right) < l + \varepsilon < 1$$

$$\Rightarrow \qquad \frac{u_n}{u_{n+1}} < \frac{n+1}{n}, \quad \forall\, n \geq m$$

If $v_n = 1/n$, the series $\sum v_n$ is divergent, and

$$\frac{u_n}{u_{n+1}} < \frac{v_n}{v_{n+1}}$$

Hence by comparison test the series $\sum u_n$ diverges.

Notes:

1. The test fails to give any definite information for $l = 1$. Consider the two series $\sum \frac{1}{n}$ and $\sum \frac{1}{n(\log n)^2}$.

   The former is divergent, while the latter is convergent but

   $$\lim_{n\to\infty} n\left(\frac{u_n}{u_{n+1}} - 1\right) = 1$$

   for both.

2. Raabe's Test is stronger that D'Alembert's Ratio Test and may succeed where Ratio-Test fails.

Example 12. Test for convergence of the series

$$\frac{\alpha}{\beta} + \frac{1+\alpha}{1+\beta} + \frac{(1+\alpha)(2+\alpha)}{(1+\beta)(2+\beta)} + \ldots$$

■ Here

$$u_n = \frac{(1+\alpha)(2+\alpha)(3+\alpha)\ldots(n-1+\alpha)}{(1+\beta)(2+\beta)(3+\beta)\ldots(n-1+\beta)}$$

$$\lim_{n\to\infty} \frac{u_{n+1}}{u_n} = \lim_{n\to\infty} \frac{n+\alpha}{n+\beta} = 1$$

Hence, the Ratio-Test fails.

Again

$$\lim_{n\to\infty} n\left(\frac{u_n}{u_{n+1}} - 1\right) = \lim_{n\to\infty} n\left(\frac{n+\beta}{n+\alpha} - 1\right) = \lim_{n\to\infty} \frac{\beta-\alpha}{1+\alpha/n} = \beta - \alpha$$

Thus by Raabe's Test, the series converges if $\beta - \alpha > 1$ or $\beta > \alpha + 1$, and diverges if $\beta < \alpha + 1$. The test fails for $\beta = \alpha + 1$.

But for $\beta = \alpha + 1$, the series becomes

$$\frac{\alpha}{\alpha+1} + \frac{1+\alpha}{2+\alpha} + \frac{1+\alpha}{3+\alpha} + \ldots = \sum \frac{1+\alpha}{n+\alpha}$$

which diverges, by comparison with $\sum 1/n$.

Example 13. Show that the series

$$\Sigma \frac{3\cdot 6\cdot 9 \ldots 3n}{7\cdot 10\cdot 13 \ldots (3n+4)} x^n, \quad x > 0$$

converges for $x \leq 1$, and diverges for $x > 1$.

■ Now

$$\frac{u_n}{u_{n+1}} = x^n \frac{3\cdot 6\cdot 9 \ldots 3n}{7\cdot 10 \ldots (3n+4)} \cdot \frac{7\cdot 10\cdot 13 \ldots (3n+4)(3n+7)}{3\cdot 6\cdot 9 \ldots 3n(3n+3)} \frac{1}{x^{n+1}}$$

$$= \frac{3n+7}{3n+3} \cdot \frac{1}{x}$$

$$\therefore \qquad \lim_{n\to\infty} \frac{u_{n+1}}{u_n} = x$$

By Ratio Test, the series converges for $x < 1$ and diverges for $x > 1$. The test fails for $x = 1$. But for $x = 1$,

$$\frac{u_n}{u_{n+1}} = \frac{3n+7}{3n+3}$$

$$\therefore \qquad \lim_{n\to\infty} n\left(\frac{u_n}{u_{n+1}} - 1\right) = \lim_{n\to\infty} \frac{4n}{3n+3} = \frac{4}{3} > 1$$

$\therefore$ By Raabe's Test, the series converges.

Thus, the series converges for $0 < x \leq 1$ and diverges for $x > 1$.

EXERCISE

Test the behaviour of the following series:

1. $\frac{1}{2} + \frac{1}{2\cdot 2^2} + \frac{1}{3\cdot 2^3} + \frac{1}{4\cdot 2^4} + \ldots$

2. $\left(\frac{1}{3}\right)^2 + \left(\frac{1\cdot 2}{3\cdot 5}\right)^2 + \left(\frac{1\cdot 2\cdot 3}{3\cdot 5\cdot 7}\right)^2 + \ldots$

3. $\frac{x}{1} + \frac{x^2}{2} + \frac{x^3}{3} + \frac{x^4}{4} + \ldots, \; x > 0$

4. $\frac{x^2}{2\sqrt{1}} + \frac{x^3}{3\sqrt{2}} + \frac{x^4}{4\sqrt{3}} + \frac{x^5}{5\sqrt{4}} + \ldots, x > 0$

5. $1 + \frac{x^2}{2^p} + \frac{x^4}{4^p} + \frac{x^6}{6^p} + \ldots$

6. $\dfrac{1^2 \cdot 2^2}{1!} + \dfrac{2^2 \cdot 3^2}{2!} + \dfrac{3^2 \cdot 4^2}{3!} + \dfrac{4^2 \cdot 5^2}{4!} + \ldots$

7. $\left(\dfrac{2^2}{1^2} - \dfrac{2}{1}\right)^{-1} + \left(\dfrac{3^3}{2^3} - \dfrac{3}{2}\right)^{-2} + \left(\dfrac{4^4}{3^4} - \dfrac{4}{3}\right)^{-3} + \cdots$

8. $\dfrac{1}{2} + \dfrac{1}{3^2} + \dfrac{1}{2^3} + \dfrac{1}{3^4} + \ldots$

Test for convergence each of the following series whose nth terms are given:

9. $\dfrac{\lfloor n}{n^n}$

10. $\dfrac{1 \cdot 2 \cdot 3 \ldots n}{7 \cdot 10 \ldots (3n+4)}$

11. $\dfrac{n^3+5}{3^n+2}$

12. $\dfrac{r^n}{n^n},\ r > 0$

13. $\dfrac{\sqrt{n}\, x^n}{\sqrt{n^2+1}},\ x > 0$

14. $\dfrac{1 \cdot 3 \cdot 5 \ldots (2n-1)}{2 \cdot 4 \cdot 6 \ldots 2n} \cdot \dfrac{1}{n}$

15. $\dfrac{1 \cdot 3 \cdot 5 \ldots (4n-3)}{2 \cdot 4 \cdot 6 \ldots (4n-2)} \cdot \dfrac{x^{2n}}{4n}$

16. $\sqrt{\dfrac{n-1}{n^3+1}}\, x^n,\ x > 0$

17. $\dfrac{2 \cdot 4 \cdot 6 \ldots (2n+2)}{3 \cdot 5 \cdot 7 \ldots (2n+3)}\, x^{n-1},\ x > 0$

ANSWERS

1. Converges
2. Converges
3. Converges for $x < 1$, diverges for $x \geq 1$
4. Converges for $x \leq 1$, diverges for $x > 1$
5. Converges for $|x| < 1$, diverges for $|x| > 1$; at $x = \pm 1$, converges for $p > 1$, diverges for $p \leq 1$
6. Convergent
7. Convergent
8. Convergent
9. Convergent
10. Convergent
11. Convergents
12. Convergent
13. Convergent for $x < 1$, divergent for $x \geq 1$
14. Convergent
15. Convergent for $|x| \leq 1$, divergent for $|x| > 1$
16. Convergent for $x < 1$, divergent for $x \geq 1$
17. Convergent for $x < 1$, divergent for $x \geq 1$.

7. LOGARITHMIC TEST

If Σu_n is a positive term series such that,

$$\lim_{n\to\infty}\left(n \log \frac{u_n}{u_{n+1}}\right) = l,$$

then the series converges for $l > 1$, and diverges for $l < 1$.

First, let $l > 1$.

Let us select $\varepsilon > 0$, such that $l - \varepsilon > 1$.

Let $l - \varepsilon = \alpha > 1$.

Since $\lim_{n\to\infty}\left(n \log \frac{u_n}{u_{n+1}}\right) = l$, therefore $\exists$ an m, such that

$$l - \varepsilon < n \log \frac{u_n}{u_{n+1}} < l + \varepsilon, \ \forall\, n \geq m$$

$$\Rightarrow \qquad n \log \frac{u_n}{u_{n+1}} > \alpha, \ \forall\, n \geq m$$

$$\Rightarrow \qquad \frac{u_n}{u_{n+1}} > e^{\alpha/n}, \ \forall\, n \geq m$$

Now since $\{(1 + 1/n)^n\}$ is a monotonic increasing sequence converging to e, therefore

$$\left(1 + \frac{1}{n}\right)^n \leq e, \ \forall\, n$$

so that we get

$$\frac{u_n}{u_{n+1}} > \left(1 + \frac{1}{n}\right)^\alpha = \frac{(n+1)^\alpha}{n^\alpha} = \frac{v_n}{v_{n+1}}, \ \forall\, n \geq m$$

where $v_n = 1/n^\alpha$.

But since for $\alpha > 1$, Σv_n converges, therefore by comparison test, Σu_n also converges.

We may similarly show that for $l < 1$, the series Σu_n diverges.

Example 14. Test for convergence of the series

$$1 + \frac{x}{1!} + \frac{2^2 x^2}{2!} + \frac{3^3 x^3}{3!} + \ldots, \text{ for } x > 0$$

- Ignoring the first term,

CHAPTER 4

$$u_n = \frac{n^n x^n}{n!}$$

Now,

$$\lim_{n\to\infty} \frac{u_{n+1}}{u_n} = \lim_{n\to\infty} \left(\frac{n+1}{n}\right)^n x = ex$$

By Ratio Test, the series converges for $ex < 1$ or $x < 1/e$, and diverges for $x > 1/e$.
For $x = 1/e$,

$$\frac{u_n}{u_{n+1}} = \left(\frac{n}{n+1}\right)^n \cdot e$$

Now,

$$\lim_{n\to\infty} \left(n \log \frac{u_n}{u_{n+1}}\right) = \lim_{n\to\infty} n \left[1 - n \log\left(1 + \frac{1}{n}\right)\right].$$

$$= \lim_{n\to\infty} n \left[1 - n\left(\frac{1}{n} - \frac{1}{2n^2} + \frac{1}{3n^3} - \dots\right)\right] = \frac{1}{2} < 1$$

Therefore by logarithmic test, the series diverges.

Hence the series converges for $x < 1/e$, and diverges for $x \geq 1/e$.

Note: Logarithmic test is generally more helpful in situations like those above, where the presence of a number like e in u_n/u_{n+1} makes the application of Raabe's Test difficult.

8. INTEGRAL TEST

Improper Integral. As preparatory to the introduction of Cauchy's Integral Test, it will help to remember that the infinite integral $\int_a^\infty u(x)\,dx$ is said to converge if $t(x) = \int_a^x u(x)\,dx$ tends to a finite limit as $x \to \infty$, otherwise the integral is said to diverge.

Further, if $u(x) \geq 0$ for all $x > a$, it can be shown geometrically or otherwise that the integral $\int_a^t u(x)\,dx$ is a monotonic increasing function of t, so that the improper integral

$$\int_a^\infty u(x)\,dx, \text{ where } u(x) \geq 0, \ \forall\ x > a$$

converges iff it is bounded above, *i.e.*, $\exists$ a positive number k such that

$$\int_a^t u(t)\,dt \leq k, \ \forall\ t \geq a$$

8.1 Cauchy's Integral Test

If u is a non-negative monotonic decreasing integrable function such that $u(n) = u_n$, for all positive integral values of n, then the series $\sum_{n=1}^{\infty} u_n$ and $\int_1^{\infty} u(x)\,dx$ converge or diverge together.

As u is monotonic decreasing, we have

$$u(n) \geq u(x) \geq u(n+1), \text{ whenever } n \leq x \leq n+1$$

Also, since u is non-negative and integrable,

$$\int_n^{n+1} u(n)\,dx \geq \int_n^{n+1} u(x)\,dx \geq \int_n^{n+1} u(n+1)\,dx$$

$$\Rightarrow \qquad u(n) \geq \int_n^{n+1} u(x)\,dx \geq u(n+1)$$

or

$$u_n \geq \int_n^{n+1} u(x)\,dx \geq u_{n+1} \qquad \text{...(1)}$$

Let us write $S_n = u_1 + u_2 + \ldots + u_n$, and $I_n = \int_1^n u(x)\,dx$, and putting $n = 1, 2, \ldots, (n-1)$ successively, and adding, we get

$$S_n - u_n \geq I_n \geq S_n - u_1$$

$$\Rightarrow \qquad 0 < u_n \leq S_n - I_n \leq u_1 \qquad \text{...(2)}$$

Let us consider the sequence $\{(S_n - I_n)\}$.

$$\begin{aligned}(S_n - I_n) - (S_{n-1} - I_{n-1}) &= S_n - S_{n-1} - (I_n - I_{n-1})\\ &= u_n - \int_{n-1}^{n} u(x)\,dx\\ &\leq 0 \qquad \text{[using (1)]}\end{aligned}$$

Therefore, the sequence $\{(S_n - I_n)\}$ is monotonic decreasing, bounded by 0 and u_1.

Hence, the sequence converges and has a limit such that

$$0 \leq \lim (S_n - I_n) \leq u_1 \qquad \text{...(3)}$$

Thus the series $\sum u_n$ converges or diverges with the integral $\int_1^{\infty} u(x)\,dx$; if convergent, the sum of the series differs from the integral by less than u_1; if divergent, the limit of $(S_n - I_n)$ still exists and lies between 0 and u_1.

Example 15. Show that the series $\sum(1/n^p)$ converges if $p > 1$, and diverges if $p \leq 1$.

- Let $u(x) = 1/x^p$, so that for $x \geq 1$, the function u is a non-negative monotonic decreasing integrable function such that

$$u_n = u(n) = \frac{1}{n^p}, \quad \forall\, n \in \mathbf{N}$$

By Integral Test, $\sum_{n=1}^{\infty} u_n$ and $\int_1^{\infty} u(x)\,dx$ converge or diverge together.

Let us now test the convergence of the infinite integral.

$$\because \qquad \int_1^X u(x)\,dx = \int_1^X \frac{1}{x^p}\,dx = \begin{cases} \dfrac{1}{1-p}\left(X^{1-p}-1\right), & \text{if } p \neq 1 \\ \log X, & \text{if } p = 1 \end{cases}$$

$$\therefore \qquad \int_1^{\infty} u(x)\,dx = \lim_{X \to \infty} \int_1^X u(x)\,dx = \begin{cases} \dfrac{1}{p-1}, & \text{if } \quad p > 1 \\ \infty, & \text{if } 0 < p \leq 1 \end{cases}$$

Thus $\int_1^{\infty} u(x)\,dx$ converges if $p > 1$, and diverges if $0 < p \leq 1$.

Hence, the infinite series $\Sigma(1/n^p)$ converges if $p > 1$, and diverges if $0 < p \leq 1$.

But, when $p < 0$, the series $\Sigma(1/n^p)$ diverges for then the nth term n^{-p} does not tend to zero as $n \to \infty$.

Hence the series $\Sigma(1/n^p)$ converges when $p > 1$, and diverges when $p \leq 1$.

Example 16. The series $\sum_{n=2}^{\infty} \dfrac{1}{n(\log n)^p}$, $p > 0$ converges for $p > 1$ and diverges for $p \leq 1$.

■ Let $u(x) = \dfrac{1}{x(\log x)^p}$, so that for $x \geq 2$, the function u is a non-negative monotonic decreasing integrable function, such that

$$u_n = u(n) = \frac{1}{n(\log n)^p}, \quad \forall\, p > 0, n \in \mathbf{N}$$

By Integral Test $\sum_{n=2}^{\infty} u_n$ and $\int_2^{\infty} u(x)\,dx$ converge or diverge together.

Let us now test the convergence of the infinite integral.

$$\because \qquad \int_2^X u(x)\,dx = \int_2^X \frac{1}{x(\log x)^p}\,dx, \quad p > 0$$

$$= \begin{cases} \dfrac{(\log X)^{1-p} - (\log 2)^{1-p}}{1-p}, & \text{if } p \neq 1, \\ \log\log X - \log\log 2, & \text{if } p = 1. \end{cases}$$

$$\therefore \qquad \int_2^{\infty} u(x)\,dx = \lim_{X\to\infty}\int_2^{X} u(x)\,dx = \begin{cases} \dfrac{(\log 2)^{1-p}}{p-1}, & \text{if } p>1 \\ \infty, & \text{if } 0<p\le 1. \end{cases}$$

Thus $\int_2^{\infty} u(x)\,dx$ converges if $p > 1$, and diverges if $0 < p \le 1$.

Hence the series $\sum\limits_{n=2}^{\infty} \dfrac{1}{n(\log n)^p}$, $p > 0$, converges if $p > 1$, and diverges if $p \le 1$.

9. GAUSS'S TEST

If $\sum u_n$ is a positive terms series such that,

$$\frac{u_n}{u_{n+1}} = \alpha + \frac{\beta}{n} + \frac{\gamma_n}{n^p},$$

where $\alpha > 0$, $p > 1$, and $\{\gamma_n\}$ is a bounded sequence, then

(*i*) *for $\alpha \ne 1$, $\sum u_n$ converges if $\alpha > 1$, and diverges if $\alpha < 1$, whatever β may be*

(*ii*) *for $\alpha = 1$, $\sum u_n$ converges if $\beta > 1$, and diverges if $\beta \le 1$.*

(*i*) When $\alpha \ne 1$,

$$\lim_{n\to\infty}\frac{u_n}{u_{n+1}} = \alpha$$

Hence by Ratio Test, the series converges if $\alpha > 1$, and diverges if $\alpha < 1$.

(*ii*) When $\alpha = 1$,

$$\lim_{n\to\infty} n\left(\frac{u_n}{u_{n+1}} - 1\right) = \beta$$

Hence by Raabe's Test, the series converges if $\beta > 1$, and diverges if $\beta < 1$.

For $\beta = 1$, we have

$$\frac{u_n}{u_{n+1}} = 1 + \frac{1}{n} + \frac{\gamma_n}{n^p}, \quad p > 1$$

Let us compare the given series with the divergent series $\sum v_n$ where $v_n = \dfrac{1}{n\log n}$.

Now,

$$\frac{u_n}{u_{n+1}} - \frac{v_n}{v_{n+1}} = 1 + \frac{1}{n} + \frac{\gamma_n}{n^p} - \frac{(n+1)\log(n+1)}{n\log n}$$

$$= \frac{\gamma_n}{n^p} - \frac{n+1}{n}\left[\frac{\log(n+1)}{\log n} - 1\right]$$

$$= \frac{1}{n^p}\left[\gamma_n - (n+1)\log\left(1+\frac{1}{n}\right)\cdot\frac{n^{p-1}}{\log n}\right]$$

But $\lim\limits_{n\to\infty}(n+1)\log\left(1+\frac{1}{n}\right) = \lim\limits_{n\to\infty}\left[\log\left(1+\frac{1}{n}\right)^n + \log\left(1+\frac{1}{n}\right)\right] = 1$

and $\lim\limits_{n\to\infty}\frac{n^{p-1}}{\log n} = \infty$, $p > 1$, and $\{\gamma_n\}$ is bounded, therefore, for sufficiently large values of n,

$\gamma_n - (n+1)\log\left(1+\frac{1}{n}\right)\cdot\frac{n^{p-1}}{\log n}$ remains negative.

Thus $\exists$ a positive integer m such that

$$\gamma_n - (n+1)\log\left(1+\frac{1}{n}\right)\cdot\frac{n^{p-1}}{\log n} < 0, \ \forall\, n \ge m$$

$$\Rightarrow \qquad \frac{u_n}{u_{n+1}} - \frac{v_n}{v_{n+1}} < 0, \ \forall\, n \ge m$$

$$\Rightarrow \qquad \frac{u_n}{u_{n+1}} < \frac{v_n}{v_{n+1}}, \ \forall\, n \ge m$$

Since $\sum v_n$ is divergent, therefore by Comparison Test, the series $\sum u_n$ is also divergent.

Remarks:

1. Gauss's Test is very useful and may be used after the failure of Raabe's Test or directly without recourse to other tests.
2. If

$$\frac{u_n}{u_{n+1}} = \alpha + \frac{\beta}{n} + \frac{\gamma}{n^2} + \frac{\delta}{n^3} + \ldots \qquad \ldots(1)$$

where $\alpha, \beta, \gamma, \ldots$ are independent of n, then we can write

$$\frac{u_n}{u_{n+1}} = \alpha + \frac{\beta}{n} + \frac{\gamma_n}{n^2}$$

where $\gamma_n = \gamma + \delta/n + \ldots$, so that $\lim \gamma_n = \gamma$, *i.e.*, $\{\gamma_n\}$ is a bounded sequence.

Thus for the application of Gauss's Test, we may expand u_n/u_{n+1} in powers of $1/n$ as in (1).

Example 17. Test the convergence of the series

$$\frac{2^2}{3^2} + \frac{2^2 \cdot 4^2}{3^2 \cdot 5^2} + \frac{2^2 \cdot 4^2 \cdot 6^2}{3^2 \cdot 5^2 \cdot 7^2} + \ldots$$

■ Now

$$u_n = \frac{2^2 \cdot 4^2 \cdot 6^2 \ldots (2n)^2}{3^2 \cdot 5^2 \cdot 7^2 \ldots (2n+1)^2}$$

$$\therefore \qquad \frac{u_n}{u_{n+1}} = \frac{(2n+3)^2}{(2n+2)^2} = \left(1 + \frac{3}{2n}\right)^2 \left(1 + \frac{1}{n}\right)^{-2}$$

$$= \left(1 + \frac{3}{n} + \frac{9}{4n^2}\right)\left(1 - \frac{2}{n} + \frac{3}{n^2} - \ldots\right)$$

$$= 1 + \frac{1}{n} - \frac{3}{4n^2} + \ldots \text{ higher powers of } \frac{1}{n}$$

so that $\alpha = 1$ and $\beta = 1$.

Hence by Gauss's Test, the series diverges.

Example 18. Test for convergence of the series

$$\Sigma \frac{1^2 \cdot 3^2 \ldots (2n-1)^2}{2^2 \cdot 4^2 \ldots (2n)^2} x^{n-1}, \quad x > 0.$$

■ Here

$$\lim_{n\to\infty} \frac{u_n}{u_{n+1}} = \lim_{n\to\infty} \frac{(2n+2)^2}{(2n+1)^2} \cdot \frac{1}{x} = \frac{1}{x}$$

Hence by Ratio Test, the series converges if $x < 1$, and diverges if $x > 1$

Now for $x = 1$,

$$\frac{u_n}{u_{n+1}} = \frac{(2n+2)^2}{(2n+1)^2}$$

$$\therefore \qquad \lim_{n\to\infty} n\left(\frac{u_n}{u_{n+1}} - 1\right) = \lim_{n\to\infty} \frac{4n^2 + 3n}{(2n+1)^2} = 1$$

Hence Raabe's Test fails.

Let us now apply Gauss's Test.

$$\frac{u_n}{u_{n+1}} = \left(1 + \frac{1}{n}\right)^2 \left(1 + \frac{1}{2n}\right)^{-2}$$

CHAPTER 4

$$= \left(1 + \frac{2}{n} + \frac{1}{n^2}\right)\left(1 - \frac{1}{n} + \frac{3}{4n^2} + \ldots\right)$$

$$= 1 + \frac{1}{n} - \frac{1}{4n^2} + \ldots \text{ higher powers of } \frac{1}{n}$$

so that by Gauss's Test, the series diverges.

Hence, the series converges for $x < 1$, and diverges for $x \geq 1$.

Note: We could get the result directly by Gauss's Test, for

$$\frac{u_n}{u_{n+1}} = \frac{1}{x}\left(\frac{2n+2}{2n+1}\right)^2 = \frac{1}{x} + \frac{1/x}{n} - \frac{1/4x}{n^2} + \ldots$$

where $\alpha = 1/x$, $\beta = 1/x$.

Example 19. Test for convergence of the hypergeometric series

$$1 + \frac{\alpha \cdot \beta}{1 \cdot \gamma}x + \frac{\alpha(\alpha+1)\,\beta(\beta+1)}{1 \cdot 2\gamma(\gamma+1)}x^2 + \frac{\alpha(\alpha+1)(\alpha+2)\,\beta(\beta+1)(\beta+2)}{1 \cdot 2 \cdot 3\,\gamma(\gamma+1)(\gamma+2)}x^3 + \ldots$$

for all positive values of x; α, β, γ being all positive.

- It is a positive term series.

Ignoring the first term, which does not affect the behaviour of the series, we have

$$u_n = \frac{\alpha(\alpha+1)\ldots(\alpha+n-1)}{1 \cdot 2 \cdot 3 \ldots \cdot n}\,\frac{\beta(\beta+1)\ldots(\beta+n-1)}{\gamma(\gamma+1)\ldots(\gamma+n-1)}\,x^n$$

so that

$$\lim_{n\to\infty} \frac{u_n}{u_{n+1}} = \lim_{n\to\infty} \frac{(n+1)(\gamma+n)}{(\alpha+n)(\beta+n)}\frac{1}{x} = \frac{1}{x}$$

Hence, by Ratio Test, the series converges if $x < 1$, and diverges if $x > 1$. For $x = 1$, we have

$$\frac{u_n}{u_{n+1}} = \frac{(n+1)(\gamma+n)}{(\alpha+n)(\beta+n)} = \frac{\left(1+\frac{1}{n}\right)\left(1+\frac{\gamma}{n}\right)}{\left(1+\frac{\alpha}{n}\right)\left(1+\frac{\beta}{n}\right)} = \frac{\left(1+\frac{\gamma+1}{n}+\frac{\gamma}{n^2}\right)}{\left(1+\frac{\alpha+\beta}{n}+\frac{\alpha\beta}{n^2}\right)}$$

$$= \left(1+\frac{\gamma+1}{n}+\frac{\gamma}{n^2}\right)\left[1-\left(\frac{\alpha+\beta}{n}+\frac{\alpha\beta}{n^2}\right)+\left(\frac{\alpha+\beta}{n}+\frac{\alpha\beta}{n^2}\right)^2+\ldots\right]$$

$$= 1 + \frac{1+\gamma-\alpha-\beta}{n} + \frac{(\alpha+\beta-\gamma)(\alpha+\beta-1)-\alpha\beta}{n^2} + \ldots$$

Hence by Gauss's Test the series converges if $1 + \gamma - \alpha - \beta > 1$ or $\gamma > \alpha + \beta$ and diverges if $1 + \gamma - \alpha - \beta \le 1$ or $\gamma \le \alpha + \beta$.

Thus for positive values of α, β, γ and x,

(*i*) for $x < 1$, the series converges,

(*ii*) for $x > 1$, the series diverges, and

(*iii*) for $x = 1$, the series converges if $\gamma > (\alpha + \beta)$ and diverges if $\gamma \le (\alpha + \beta)$.

EXERCISE

Test the convergence of the series:

1. $1 + \dfrac{2x}{2!} + \dfrac{3^2 \cdot x^2}{3!} + \dfrac{4^3 \cdot x^3}{4!} + \ldots$, for $x > 0$

2. $\dfrac{x}{1} + \dfrac{1}{2}\dfrac{x^3}{3} + \dfrac{1 \cdot 3}{2 \cdot 4}\dfrac{x^5}{5} + \dfrac{1 \cdot 3 \cdot 5}{2 \cdot 4 \cdot 6}\dfrac{x^7}{7} + \ldots$, for $x > 0$

3. $\dfrac{a + x}{1!} + \dfrac{(a + 2x)^2}{2!} + \dfrac{(a + 3x)^3}{3!} + \ldots$, for $a, x > 0$.

4. $\dfrac{1^2}{2^2} + \dfrac{1^2 \cdot 3^2}{2^2 \cdot 4^2} + \dfrac{1^2 \cdot 3^2 \cdot 5^2}{2^2 \cdot 4^2 \cdot 6^2} + \ldots$

5. Apply Cauchy's Integral Test to test the convergence of the series.

   (*i*) $\sum_{n=1}^{\infty} \dfrac{1}{n^2 + n}$, (*ii*) $\sum_{n=1}^{\infty} \dfrac{1}{n^2 + 1}$

6. Test the convergence of $\sum_{n=3}^{\infty} \dfrac{1}{n \log n \, (\log \log n)^p}$, $p > 0$

   Test for convergence the series whose nth term is

7. $\dfrac{(n!)^2}{(2n)!} x^n$, $x > 0$,

8. $\dfrac{n!}{(n + 1)^n} x^n$, $x > 0$.

9. Prove that $1 + \dfrac{\alpha}{\beta} + \dfrac{\alpha(\alpha + 1)}{\beta(\beta + 1)} + \dfrac{\alpha(\alpha + 1)(\alpha + 2)}{\beta(\beta + 1)(\beta + 2)} + \ldots$, where α and β are positive, converges if $\beta > \alpha + 1$, and diverges if $\beta \le \alpha + 1$.

ANSWERS

1. Converges if $x < 1/e$, diverges if $x \geq 1/e$.
2. Converges if $|x| \leq 1$, diverges if $|x| > 1$.
3. Converges for $x < 1/e$, diverges for $x \geq 1/e$.
4. Diverges
5. (*i*) Convergent (*ii*) Convergent.
6. Converges for $p > 1$, diverges for $p \leq 1$.
7. Converges for $x < 4$, diverges for $x \geq 4$.
8. Converges for $0 < x < e$, diverges for $x \geq e$.

10. SERIES WITH ARBITRARY TERMS

So far we have considered series with positive terms only. We shall now discuss series with terms having any sign whatsoever.

10.1 Alternating Series

A series whose terms are alternatively positive and negative, *e.g.*,

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \ldots$$

is called an *alternating series*.

Leibnitz Test. *If the alternating series*

$$u_1 - u_2 + u_3 - u_4 + \ldots, \left(u_n > 0, \forall\, n\right)$$

is such that

(*i*) $u_{n+1} \leq u_n, \ \forall\, n$, *and*

(*ii*) $\lim\limits_{n \to \infty} u_n = 0$,

then the series converges.

Let $S_n = u_1 - u_2 + u_3 - u_4 + \ldots + (-1)^{n-1}\, u_n$

Now for all n,

$$S_{2n+2} - S_{2n} = u_{2n+1} - u_{2n+2} \geq 0$$

$\Rightarrow$ $$S_{2n+2} \geq S_{2n}$$

$\Rightarrow$ $\{S_{2n}\}$ is a monotonic increasing sequence.

Again

$$S_{2n} = u_1 - u_2 + u_3 - \ldots + u_{2n-1} - u_{2n}$$
$$= u_1 - (u_2 - u_3) - (u_4 - u_5) - \ldots - (u_{2n-2} - u_{2n-1}) - u_{2n}$$

But since $u_{n+1} \leq u_n$, for all n, therefore, each bracket on the right is positive and hence

$$S_{2n} < u_1, \ \forall\, n$$

Thus, the monotonic increasing sequence $\{S_{2n}\}$ is bounded above and is consequently convergent.

Let $\lim\limits_{n\to\infty} S_{2n} = S$.

We shall now show that the sequence $\{S_{2n+1}\}$ also converges to the same limit S.

Now

$$S_{2n+1} = S_{2n} + u_{2n+1}$$

$$\therefore \qquad \lim_{n\to\infty} S_{2n+1} = \lim_{n\to\infty} S_{2n} + \lim_{n\to\infty} u_{2n+1}$$

But by condition (*ii*),

$$\lim_{n\to\infty} u_{2n+1} = 0$$

$$\therefore \qquad \lim_{n\to\infty} S_{2n+1} = \lim_{n\to\infty} S_{2n} = S$$

Thus, the sequences $\{S_{2n}\}$ and $\{S_{2n+1}\}$ both converge to the same limit S. We shall now show that the sequence $\{S_n\}$ also converges to S.

Let $\varepsilon > 0$ be given.

Since the sequences $\{S_{2n}\}$ and $\{S_{2n+1}\}$ both converge to S, therefore $\exists$ positive integers m_1, m_2, respectively, such that

$$|S_{2n} - S| < \varepsilon, \ \forall\, n \geq m_1 \qquad \text{...(1)}$$

and

$$|S_{2n+1} - S| < \varepsilon, \ \forall\, n \geq m_2 \qquad \text{...(2)}$$

Thus from (1) and (2), we have

$$|S_n - S| < \varepsilon, \ \forall\, n \geq \max(2m_1, 2m_2 + 1)$$

$\Rightarrow$ $\{S_n\}$ converges to S

$\Rightarrow$ The series $\sum (-1)^{n-1} u_n$ converges.

Example 20. Show that the series $\dfrac{1}{1^p} - \dfrac{1}{2^p} + \dfrac{1}{3^p} - \dfrac{1}{4^p} + \ldots$ converges for $p > 0$.

■ Let $u_n = 1/n^p$.

Here

$$u_{n+1} \leq u_n, \ \forall\, n$$

and

$$\lim_{n\to\infty} u_n = \lim_{n\to\infty} \frac{1}{n^p} = 0$$

Hence by Leibnitz test, the alternating series $\sum \dfrac{(-1)^{n-1}}{n^p}$ converges.

CHAPTER 4

10.2 Absolute Convergence

A series Σu_n is said to be *absolutely convergent* if the series Σu_n obtained on taking every term of the given series with a positive sign is convergent, *i.e.*, if the series $\Sigma|u_n|$ is convergent.

Conditional Convergence. A series which is convergent but is not absolutely convergent is called a *conditionally convergent* series.

ILLUSTRATIONS

1. The series
$$1-\frac{1}{2^3}+\frac{1}{3^3}-\frac{1}{4^3}+\ldots$$
is absolutely convergent because the series
$$1+\frac{1}{2^3}+\frac{1}{3^3}+\frac{1}{4^3}+\ldots$$
obtained on taking every term of the given series with a positive sign, is convergent.
2. The series
$$1-\frac{1}{2}+\frac{1}{3}-\frac{1}{4}+\ldots$$
is convergent by Leibnitz test, but the series
$$1+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\ldots$$
obtained on taking every term with a positive sign is divergent. Thus, the given series is conditionally convergent.

Theorem 10. *Every absolutely convergent series is convergent.*

Let Σu_n be absolutely convergent, so that $\Sigma|u_n|$ is convergent.

Hence, for any $\varepsilon > 0$, by Cauchy's General Principle of convergence, $\exists$ a positive integer m such that
$$|u_{n+1}|+|u_{n+2}|+\ldots+|u_{n+p}|<\varepsilon, \quad \forall\, n\geq m \wedge p\geq 1$$
Also for all n and $p > 1$,
$$|u_{n+1}+u_{n+2}+\ldots+u_{n+p}|\leq|u_{n+1}|+|u_{n+2}|+\ldots+|u_{n+p}|<\varepsilon, \quad \forall\, n\geq m \wedge p\geq 1.$$
Hence, by Cauchy's General Principle of convergence the series Σu_n converges.

Aliter. Let Σu_n be absolutely convergent, so that $\Sigma|u_n|$ is convergent.

Let $a_n=\begin{cases}u_n, \text{ if } u_n\geq 0\\ 0, \text{ if } u_n<0\end{cases}$ and $b_n=\begin{cases}-u_n, \text{ if } u_n<0\\ 0, \quad \text{if } u_n\geq 0.\end{cases}$

Then clearly,
$$a_n\geq 0,\, b_n\geq 0, \qquad \ldots(1)$$

$$u_n = a_n - b_n \qquad ...(2)$$

and

$$|u_n| = a_n + b_n \qquad ...(3)$$

From (1) and (3), it follows that

$$a_n \le |u_n|, b_n \le |u_n|$$

Since $\Sigma|u_n|$ is convergent, therefore, by Comparison Test, both Σa_n and Σb_n are convergent.

Hence, by Theorem 5, $\Sigma(a_n - b_n)$ is convergent.

Hence from (2), it follows that Σu_n is convergent.

Remarks:

1. The divergence of $\Sigma|u_n|$ does not imply the divergence of Σu_n.

   For example, if $u_n = \frac{(-1)^{n-1}}{n}$, we have seen above that $\Sigma|u_n|$ is divergent, whereas Σu_n is convergent.

2. The very great significance of the concept of Absolute Convergence is that the convergence of absolutely convergent series is much more easy to recognise than that of conditionally convergent series—usually by comparison with series of positive terms. In fact all the tests for positive term series become available for the purpose. But this significance becomes more visible in the discussion of rearrangement of series—so much so that we may operate on absolutely convergent series, precisely as we operate on sums of a finite number of terms, whereas in the case of conditionally convergent series this in general is not possible.

Example 21. Show that for any fixed value of x, the series $\sum_{n=1}^{\infty} \frac{\sin nx}{n^2}$ is convergent.

■ Let $u_n = \frac{\sin nx}{n^2}$, so that $|u_n| = \frac{|\sin nx|}{n^2}$

Now $\left|\frac{\sin nx}{n^2}\right| \le \frac{1}{n^2}$, $\forall\ n$ and $\Sigma\frac{1}{n^2}$ converges.

Hence, by Comparison Test, the series $\Sigma\left|\frac{\sin nx}{n^2}\right|$ converges.

Since every absolutely convergent series is convergent, therefore $\Sigma\frac{\sin nx}{n^2}$ is convergent.

Example 22. Show that the series

$$x + \frac{x^2}{2!} + \frac{x^3}{3!} + ...$$

converges absolutely for all values of x.

■ Let $u_n = \frac{x^n}{n!}$.

Now, $\lim\limits_{n\to\infty} \dfrac{|u_n|}{|u_{n+1}|} = \lim\limits_{n\to\infty} \dfrac{n+1}{|x|} \to \infty$ except when $x = 0$.

So by Ratio Test, the series converges absolutely for all x except possibly zero.

But for $x = 0$ the series evidently converges absolutely.

Hence, the series converges absolutely for all values of x.

Note: Since for a convergent series Σu_n, $\lim\limits_{n\to\infty} u_n = 0$,

$$\therefore \qquad \lim_{n\to\infty} \frac{x^n}{n!} = 0$$

A useful result.

Ex. If $b > 0$, then show that the series

$$x + \frac{a-b}{\lfloor 2} x^2 + \frac{(a-b)(a-2b)}{\lfloor 3} x^3 + \ldots \text{ converges absolutely for } |x| < b^{-1}.$$

Example 23. Show that

$$\lim_{n\to\infty} \frac{m(m-1)\ldots(m-n+1)}{(n-1)!} x^n = 0,$$

where $|x| < 1$ and m is any real number.

■ Consider the series Σu_n, where

$$u_n = \frac{m(m-1)\ldots(m-n+1)}{(n-1)!} x^n$$

Now

$$\lim_{n\to\infty} \frac{|u_n|}{|u_{n+1}|} = \lim_{n\to\infty} \left|\frac{n}{m-n}\right| \cdot \frac{1}{|x|}$$

$$= \lim_{n\to\infty} \left|\frac{1}{\frac{m}{n} - 1}\right| \cdot \frac{1}{|x|} = \frac{1}{|x|}$$

Hence, the series Σu_n converges absolutely for $|x| < 1$

$\Rightarrow$ The series Σu_n converges for $|x| < 1$

$\Rightarrow$ $\lim\limits_{n\to\infty} u_n = 0$, for $|x| < 1$

i.e., $$\lim_{n\to\infty} \frac{m(m-1)\ldots(m-n+1)}{(n-1)!} x^n = 0, \text{ if } |x| < 1$$

Note: The results of Examples 22 and 23 are very useful.

Also see Examples 12 and 13 of Ch. 3.

EXERCISE

1. Show that the following series are convergent:

   (i) $1-\frac{1}{2}+\frac{1}{3}-\frac{1}{4}+...$

   (ii) $\frac{1}{1^2}-\frac{1}{2^2}+\frac{1}{3^2}-\frac{1}{4^2}+...$

   (iii) $1-\frac{1}{2\sqrt{2}}+\frac{1}{3\sqrt{3}}-\frac{1}{4\sqrt{4}}+...$

   (iv) $\frac{\log 2}{2^2}-\frac{\log 3}{3^2}+\frac{\log 4}{4^2}-...$

2. Prove that the following series are absolutely convergent:

   (i) $1+\frac{x}{1!}+\frac{x^2}{2!}+...$

   (ii) $1-\frac{x^2}{2!}+\frac{x^4}{4!}-\frac{x^6}{6!}+...$

   (iii) $x-\frac{x^3}{3!}+\frac{x^5}{5!}-...$

3. Show that $\lim\limits_{n\to\infty}\frac{n^r}{x^n}=0$, if $x>1$.

4. Show that the following series are conditionally convergent:

   (i) $\Sigma\frac{(-1)^{n+1}}{\sqrt{n}}$

   (ii) $\Sigma\frac{(-1)^{n+1}}{3n-2}$

5. Show that the series $\Sigma\frac{(-1)^{n+1}}{n^p}$ is absolutely convergent for $p>1$, but conditionally convergent for $0<p\le 1$.

6. Show that the following series are absolutely convergent:

   (i) $\Sigma(-1)^{n-1}\left\{\frac{1}{n^2}+\frac{1}{(n+1)^2}\right\}$

   (ii) $\Sigma(-1)^{n-1}\left\{\frac{1}{n^{5/2}}+\frac{1}{(n+1)^{5/2}}\right\}$

(iii) $\Sigma(-1)^n \dfrac{n+2}{2^n+5}$

7. Show that the series $\Sigma\left(\dfrac{1}{n}+\dfrac{(-1)^{n-1}}{\sqrt{n}}\right)$ is divergent.

8. Show that the series $1-\dfrac{1}{3\cdot 4}+\dfrac{1}{5\cdot 4^2}-\dfrac{1}{7\cdot 4^3}+\ldots$ converges.

9. Use Cauchy's Integral Test to show that the following series converge:

(i) $\sum\limits_{n=0}^{\infty}\left(\dfrac{1+n}{1+n^2}\right)^2$

(ii) $\sum\limits_{n=2}^{\infty}\dfrac{1}{\sqrt{n}}\log\dfrac{n+1}{n-1}$

10. Show that the following series are absolutely convergent:

(i) $\Sigma\dfrac{\sin n\alpha}{n^2}$

(ii) $\Sigma(-1)^{n+1}\dfrac{n^3}{2^n}$

(iii) $\dfrac{1}{2}-\dfrac{1}{2}\cdot\dfrac{1}{2^2}+\dfrac{1}{3}\cdot\dfrac{1}{2^3}-\dfrac{1}{4}\cdot\dfrac{1}{2^4}+\ldots$

11. Show that the following series are conditionally convergent:

(i) $\sum\limits_{n=1}^{\infty}\dfrac{(-1)^{n+1}}{\log(n+1)}$

(ii) $\sum\limits_{n=1}^{\infty}\dfrac{(-1)^n}{n-\log n}$

12. Establish the divergence of the series $2-\dfrac{3}{2}+\dfrac{4}{3}-\dfrac{5}{4}+\ldots$

13. Show that if Σa_n^2 and Σb_n^2 are convergent infinite series, then $\Sigma a_n b_n$ is an absolutely convergent series.

14. Show that if the series Σa_n is absolutely convergent, then the series $\Sigma\dfrac{n+1}{n}a_n$ is also absolutely convergent.

10.3 Tests for Series of Arbitrary Terms

We now consider arbitrary term series which are convergent (but not necessarily absolutely) and obtain tests for their convergence. We first prove an important lemma, due to Abel.

Lemma. *If b_n is a positive, monotonic decreasing function and if A_n is bounded, then the series $\Sigma A_n(b_n-b_{n+1})$ is absolutely convergent.*

Since A_n is bounded, therefore $\exists$ a positive number k such that

$$|A_n| \le k, \quad \forall n$$

Thus

$$\sum_{n=1}^{m} \left| A_n\left(b_n - b_{n+1}\right)\right| = \sum_{n=1}^{m} |A_n| \left(b_n - b_{n+1}\right) \quad \left[\because\ b_n - b_{n+1} \ge 0\right]$$

$$\le k \sum_{n=1}^{m} \left(b_n - b_{n+1}\right) \quad = k\left(b_1 - b_{m+1}\right) < kb_1.$$

Thus the sequence of partial sums of the positive term series $\Sigma \left| A_n(b_n - b_{n+1})\right|$ is bounded above by kb_1, so that the series $\Sigma \left| A_n(b_n - b_{n+1})\right|$ converges, *i.e.*, the series $\Sigma A_n(b_n - b_{n+1})$ converges absolutely.

Note: The lemma may be restated as follows:

If $\{b_n\}$ is positive, monotonic decreasing sequence and if $\{A_n\}$ is a bounded sequence, then the series $\Sigma A_n(b_n - b_{n+1})$ is absolutely convergent.

10.4 Abel's Test

If b_n is a positive monotonic decreasing function and if Σu_n is a convergent series, then the series $\Sigma u_n b_n$ is also convergent.

Let $v_n = u_n b_n$ and $S_n = \sum_{r=1}^{n} u_r, V_n = \sum_{r=1}^{n} v_r$ be nth partial sums. Then

$$V_n = u_1 b_1 + u_2 b_2 + \ldots + u_n b_n$$

$$= S_1 b_1 + (S_2 - S_1)b_2 + \ldots + \left(S_n - S_{n-1}\right)b_n$$

$$= S_1\left(b_1 - b_2\right) + S_2\left(b_2 - b_3\right) + \ldots + S_{n-1}\left(b_{n-1} - b_n\right) + S_n b_n$$

$$= \sum_{r=1}^{n-1} S_r\left(b_r - b_{r+1}\right) + S_n b_n \qquad \ldots(1)$$

Since the series Σu_n is convergent, therefore the sequence $\{S_n\}$ is also convergent and hence bounded. Also b_n is a positive and monotonic decreasing function. Therefore, by the above Lemma, the series $\Sigma S_n(b_n - b_{n+1})$ and hence the partial sum $\sum_{r=1}^{n-1} S_r(b_r - b_{r+1})$ tend to a finite limits as $n \to \infty$.

Also, since $\{b_n\}$ is monotonic decreasing and bounded below by zero, therefore $\{b_n\}$ is convergent and so b_n tends to a finite limit as $n \to \infty$. Hence, $S_n b_n$ tends to finite a limit as $n \to \infty$.

Using the above results we find from (1) that V_n tends to a finite limit as $n \to \infty$, *i.e.*, the sequence $\{V_n\}$ of partial sums of Σv_n converges. Consequently the series Σv_n or $\Sigma u_n b_n$ converges.

Corollary. *A convergent series $\sum u_n$ (which need not converge absolutely) remains convergent if its terms are each multiplied by a factor a_n, provided that the sequence $\{a_n\}$ is bounded and monotonic.*

Under the given conditions, $\{a_n\}$ converges to a limit a, say. Let us write $b_n = a - a_n$, when $\{a_n\}$ is an increasing sequence, and $b_n = a_n - a$ when $\{a_n\}$ is decreasing. Then it is clear that the sequence $\{b_n\}$ monotonically decreases to the limit zero. With this function b_n, we deduce as above that the series $\sum u_n b_n$ converges.

Also, since $\sum u_n$ and hence $\sum a u_n$ converges, the convergence of $\sum u_n a_n$ follows.

10.5 Dirichlet's Test

If b_n is a positive, monotonic decreasing function with limit zero, and if, for the series $\sum u_n$, the sequence $\{S_n\}$ of partial sums is bounded, then the series $\sum u_n b_n$ is convergent.

Using the notation of § 10.4 we get as before

$$V_n = \sum_{r=1}^{n-1} S_r (b_r - b_{r+1}) + S_n b_n \qquad \text{...(1)}$$

Since S_n is bounded and b_n is positive and monotonic decreasing, therefore, by the above Lemma, $\sum\limits_{r=1}^{n-1} S_r (b_r - b_{r+1})$ tends to a finite limit as $n \to \infty$.

Also since $b_n \to 0$ as $n \to \infty$ and since S_n is bounded, therefore $S_n b_n \to 0$ as $n \to \infty$.

Using the above results, we find from (1) that V_n tends to a finite limit as $n \to \infty$ and hence the series $\sum v_n (= \sum u_n b_n)$ converges.

The case $u_n = (-1)^{n-1}$ of the above theorem is of considerable importance.

Corollary. ***Leibnitz test is a particular case of Dirichlet's test.***

Since the sequence of partial sums of the series $\sum (-1)^{n-1}$ is bounded (for, $S_n = 0$, if n is even and $S_n = 1$, if n is odd) therefore by taking $u_n = (-1)^{n-1}$, $\sum u_n b_n$ reduces to $b_1 - b_2 + b_3 - b_4 + \ldots$. Thus we obtain, "If b_n is positive and monotonic decreasing to the limit zero, then the alternating series $b_1 - b_2 + b_3 - b_4 + \ldots$ is convergent."

Example 24. Show that the series $0 - \frac{1}{2} + \frac{1}{2^2} - \frac{1}{3} + \frac{2}{3^2} - \frac{1}{4} + \frac{3}{4^2} - \ldots$ converges.

■ The given series can be considered to have arisen as a result of multiplication of the terms of the series

$$1 - 1 + \frac{1}{2} - \frac{1}{2} + \frac{1}{3} - \frac{1}{3} + \frac{1}{4} - \frac{1}{4} + \ldots \qquad \text{...(1)}$$

by the terms of the sequence

$$0, \frac{1}{2}, \frac{1}{2}, \frac{2}{3}, \frac{2}{3}, \frac{3}{4}, \frac{3}{4} \ldots \qquad \text{...(2)}$$

Since the series (1) is convergent and the sequence (2) is monotonic and bounded, therefore by Abel's Test, the given series converges.

Example 25. Test the convergence of the series

$$\sum \frac{(n^3+1)^{1/3}-n}{\log n}$$

■ Let $u_n = \{(n^3+1)^{1/3} - n\}$, and $b_n = \dfrac{1}{\log n}$.

Then the given series can be written as $\sum b_n u_n$.

Since $\sum u_n$ converges and $\{b_n\}$ is a positive monotonic decreasing sequence, therefore, by Abel's Test, the given series converges.

Example 26. Show that the series $1 - \dfrac{1}{3\cdot 2^2} + \dfrac{1}{5\cdot 3^2} - \dfrac{1}{7\cdot 4^2} + \ldots$ is convergent.

■ Let $\sum u_n = \dfrac{1}{1^2} - \dfrac{1}{2^2} + \dfrac{1}{3^2} - \dfrac{1}{4^2} + \ldots$, and $b_n = \dfrac{1}{2n-1}$.

Then the given series can be written as $\sum b_n u_n$.

Since $\sum u_n$ converges and b_n is positive and monotonic decreasing, therefore by Abel's Test, the given series converges.

11. REARRANGEMENT OF TERMS

It is a well known fact that a finite sum keeps the same value, no matter how the terms of the sum are arranged. This property, however, is by no means universally true for infinite series. For example, the series

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + \ldots$$

converges to a sum, say S. On rearranging the terms so that each positive term is followed by two negative terms, the series

$$1 - \frac{1}{2} - \frac{1}{4} + \frac{1}{3} - \frac{1}{6} - \frac{1}{8} + \frac{1}{5} - \frac{1}{10} - \frac{1}{12} + \ldots$$

converges to the sum $\dfrac{1}{2}S$.

Another rearrangement gives the series (with sign changed)

$$\frac{1}{2} + \frac{1}{4} + \frac{1}{6} + \ldots + \frac{1}{2^8} - 1 + \frac{1}{2^8+2} + \ldots + \frac{1}{2^{16}} - \frac{1}{3} + \ldots$$

which is divergent.

Consequently, the *rearrangement* (or equally well *dearrangement*) or changing the order of the terms in a series may not only alter the sum of the series but may change its nature all together. So we naturally ask *under what conditions may we rearrange the terms of the series without altering its value*?

A series $\sum v_n$ is said to arise from a series $\sum u_n$ by a *rearrangement* of terms, if there exists a one-to-one correspondence between the terms of the two series, so that every term in the first series occupies

CHAPTER 4

a perfectly definite place in the second series, and conversely. Thus, corresponding to any number of terms (say n) in the first series, we can find a number p such that p terms of the second series contain all the n terms (and some others) of the first; and conversely.

11.1 Theorem 11. *Dirichlet's theorem.* *A series obtained from an absolutely convergent series by a rearrangement of terms converges absolutely and has the same sum as the original series.*

We shall prove the theorem in two parts, first for the positive term series, and then for series of arbitrary terms.

(*i*) Let $\sum u_n$ be a given series of positive terms which converges to a sum, say S. Let $\sum v_n$ be the rearranged series. Let S_m and σ_m respectively denote the partial sums of the series $\sum u_n$ and $\sum v_n$.

As σ_m consists of m terms of the series $\sum v_n$, we can find a number p such that p terms of the series $\sum u_n$ contain all the m terms (and some more) of the former and since we are dealing with positive term series, therefore $\sigma_m \le S_p \le S$.

Thus the sequence $\{\sigma_m\}$ of partial sums of the series $\sum v_n$ of positive terms is bounded above by S. Therefore, the sequence $\{\sigma_m\}$ and consequently the series $\sum v_n$ converges to a limit, say σ, where $\sigma \le S$.

Considering $\sum u_n$ as rearrangement of $\sum v_n$, we can similarly prove that $S \le \sigma$.

Hence, $\sigma = S$, *i.e.*, the two series converge to the same sum.

(*ii*) Let $\sum u_n$ be an absolutely convergent series of arbitrary terms and $\sum v_n$, the rearrangement of $\sum u_n$.

Let

$$a_n = \begin{cases} u_n, & \text{if } u_n \ge 0 \\ 0, & \text{if } u_n < 0 \end{cases}$$

$$b_n = \begin{cases} -u_n, & \text{if } u_n < 0 \\ 0, & \text{if } u_n \ge 0 \end{cases}$$

$$a'_n = \begin{cases} v_n, & \text{if } v_n \ge 0 \\ 0, & \text{if } v_n < 0 \end{cases}$$

$$b'_n = \begin{cases} -v_n, & \text{if } v_n < 0 \\ 0, & \text{if } v_n \ge 0 \end{cases}$$

Thus clearly a_n, b_n, a'_n, b'_n are non-negative, and

$$\left.\begin{aligned} u_n = a_n - b_n, |u_n| = a_n + b_n \\ v_n = a'_n - b'_n, |v_n| = a'_n + b'_n \end{aligned}\right\} \quad \text{...(1)}$$

$$\Rightarrow \quad \left.\begin{aligned} a_n = \frac{1}{2}(|u_n| + u_n) \\ b_n = \frac{1}{2}(|u_n| - u_n) \end{aligned}\right\} \quad \text{...(2)}$$

and

$$\left.\begin{aligned} a_n' &= \frac{1}{2}\left(|v_n| + v_n\right) \\ b_n' &= \frac{1}{2}\left(|v_n| - v_n\right) \end{aligned}\right\} \qquad \ldots(3)$$

Since Σu_n is absolutely convergent, therefore from (2), Σa_n and Σb_n are also convergent. (Ref. Theorem 5). Again, since Σa_n and Σb_n are convergent series of non-negative terms, $\Sigma a_n'$ and $\Sigma b_n'$ are respectively their rearrangements, therefore by what has been proved above, $\Sigma a_n'$ and $\Sigma b_n'$ are also convergent. Also if a, b, a', b', denote the sum of the series Σa_n, Σb_n, $\Sigma a_n'$, $\Sigma b_n'$ respectively, then $a = a'$ and $b = b'$.

From (1), it follows at once that Σv_n and $\Sigma |v_n|$ are convergent and

$$\sum_{n=1}^{\infty} v_n = a' - b' = a - b = \sum_{n=1}^{\infty} u_n$$

$$\sum_{n=1}^{\infty} |v_n| = a' + b' = a + b = \sum_{n=1}^{\infty} |u_n|$$

Hence, the rearranged series Σv_n converges absolutely to the same sum as Σu_n.

Remarks:

1. For a positive term series the theorem may be stated as follows:

   A series of positive terms, if convergent, has a sum independent of the order of its terms, but if divergent, it remains divergent however its terms are rearranged.

   For the *divergent* case, one may argue as follows:

   If Σu_n is divergent, Σv_n cannot converge; for the foregoing argument shows that if Σv_n converges, Σu_n (regarded as a rearrangement of Σv_n) must also converge. Consequently Σv_n is divergent.
2. The above theorem (along with Theorem 5) shows that the brackets may be inserted or removed, or terms be picked and placed at random without changing the behaviour of a positive term series or an absolutely convergent series. In fact they behave exactly like finite sums.
3. An absolutely convergent series, because it remains convergent, with unaltered sum under every rearrangement of terms, is also called *unconditionally convergent*.

CHAPTER 4

11.2 We shall now prove a theorem which though of no practical importance, is of considerable theoretical interest.

Theorem 12. ***Riemann's theorem.*** *By an appropriate rearrangement of terms, a conditionally convergent series Σu_n can be made*

(*i*) *to converge to any number* σ, or

(*ii*) *to diverge to* $+\infty$, *or*

(*iii*) *to diverge to* $-\infty$, *or*

(*iv*) *to oscillate finitely, or*

(*v*) *to oscillate infinitely.*

Let

$$a_n = \begin{cases} u_n, & \text{if } u_n \geq 0 \\ 0, & \text{if } u_n < 0 \end{cases} \quad \text{and} \quad b_n = \begin{cases} -u_n, & \text{if } u_n < 0 \\ 0, & \text{if } u_n \geq 0. \end{cases}$$

Then clearly a_n, b_n are non-negative, and

$$u_n = a_n - b_n, \; |u_n| = a_n + b_n \qquad \text{...(1)}$$

Since $\sum u_n$ is conditionally convergent, therefore $\sum |u_n|$ diverges and hence from (1) at least one of the series $\sum a_n, \sum b_n$ diverges.

Again, since $\sum u_n$ converges, therefore it follows from (1) that the two series $\sum a_n, \sum b_n$ either both converge or both diverge (being non-negative term series, they cannot oscillate). Thus we infer that $\sum a_n$ and $\sum b_n$ both diverge.

Also $a_n \to 0, b_n \to 0$ as $n \to \infty$ $(\because\ u_n \to 0)$.

(i) We shall first show that a rearrangement $\sum v_n$, of $\sum u_n$ can be found which converges to any number, σ.

Let n_1 be the least number of terms of $\sum a_n$, whose sum

$$a_1 + a_2 + a_3 + \ldots + a_{n_1} > \sigma$$

Let m_1 be the least number of terms of $\sum b_n$, such that

$$a_1 + a_2 + a_3 + \ldots + a_{n_1} - b_1 - b_2 - \ldots - b_{m_1} < \sigma$$

Again, let n_2 be the least number of the next (terms following a_{n_1}) of $\sum a_n$, such that

$$a_1 + a_2 + \ldots + a_{n_1} - b_1 - b_2 - \ldots - b_{m_1} + a_{n_1+1} + a_{n_1+2} + \ldots + a_{n_1+n_2} > \sigma$$

Let m_2 be the least number of the next terms of $\sum b_n$ such that when $(-b_{m_1+1} - b_{m_1+2} - \ldots - b_{m_1+m_2})$ is added to the above sum, makes it less than σ. The process may be continued indefinitely. The process indicated above is always possible because of the divergence of the two series $\sum a_n, \sum b_n$.

Let $\sum v_n$ be the rearranged series and $\{\sigma_n\}$ its sequence of partial sums.

Clearly

$$\sigma_{n_1} > \sigma, \sigma_{n_1+m_1} < \sigma, \sigma_{n_1+m_1+n_2} > \sigma, \ldots$$

Therefore, it can be easily shown that the sequence $\{\sigma_n\}$ converges to σ.

$\Rightarrow$ The rearrangement $\sum v_n$ converges to σ.

(ii) We shall now show that a suitable rearrangement of $\sum u_n$ can be found which diverges to $+\infty$.

Let us consider the rearrangement

$$a_1 + a_2 + \ldots + a_m - b_1 + a_{m_1+1} + a_{m_1+2} + \ldots + a_{m_2} - b_2 + a_{m_2+1} + \ldots,$$

in which a group of positive terms is followed by a single negative term.

This is certainly a rearrangement of $\sum u_n$ and let us denote it by $\sum v_n$ and its partial sum by S_n.

Now, since the series Σa_n is divergent, its partial sums are therefore unbounded.

Let us first choose m_1 so large that

$$a_1 + a_2 + a_3 + \ldots + a_{m_1} > 1 + b_1$$

then $m_2 > m_1$ so large that

$$a_1 + a_2 + \ldots + a_{m_1} + a_{m_1+1} + \ldots + a_{m_2} > 2 + b_1 + b_2$$

and generally, $m_n > m_{n-1}$, so large that

$$a_1 + a_2 + \ldots + a_{m_n} > n + b_1 + b_2 + \ldots + b_n,$$

where $n = 1, 2, 3, \ldots$

Now, since each of the partial sum $S_{m_1+1}, S_{m_2+2} \ldots$ of Σv_n whose last term is a negative term '$-b_n$' is greater than n ($n = 1, 2, 3, \ldots$), therefore these partial sums are unbounded above and consequently the series Σv_n diverges to $+\infty$.

(*iii*) By considering the rearrangement

$$-b_1 - b_2 - \ldots - b_{m_1} + a_1 - b_{m_1+1} - b_{m_1+2} - \ldots - b_{m_2} + a_2 - b_{m_2+1} - \ldots$$

it can be shown, as before, that the rearrangement diverges to $-\infty$.

Other cases may similarly be proved by considering suitable rearrangements of the given series.

Remark: As proved earlier the absolutely convergent series remain convergent, with unaltered sum, without any condition on rearrangement of terms but it is not so in the case of non-absolutely convergent series (*i.e.*, series which converge but not absolutely). Such series change their behaviour by change in the order of the terms. This is precisely the reason why such series are called *conditionally convergent*.

Example 27. Criticise the following paradox.

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + \ldots$$

■ The given series

$$= 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \frac{1}{6} + \ldots - 2\left(\frac{1}{2} + \frac{1}{4} + \frac{1}{6} + \ldots\right)$$

$$= 1 + \frac{1}{2} + \frac{1}{3} + \ldots - \left(1 + \frac{1}{2} + \frac{1}{3} + \ldots\right)$$

$$= 0$$

Hence, the series converges to zero.

The given series is conditionally convergent and hence can be made to converge to any limit by a rearrangement of the terms.

EXERCISE

1. Assuming the convergence of Σu_n, show that the following series are convergent:

(*i*) $\Sigma \dfrac{u_n}{n}$, (*ii*) $\sum_{n=2}^{\infty} \dfrac{u_n}{\log n}$, (*iii*) $\Sigma \dfrac{u_n}{\log_a n}$,

(iv) $\Sigma\left(\frac{n+1}{n}\right)u_n$, (v) $\Sigma n^{1/n} u_n$, (vi) $\Sigma\left(1+\frac{1}{n}\right)^n u_n$.

2. Examine the convergence of the series:

(i) $\Sigma \frac{\cos n\theta}{n^\alpha}$, (ii) $\Sigma \frac{\sin n\theta}{n^\alpha}$

[*Hint*:

$$S_n = \sum_{r=1}^{n} \cos r\theta = \frac{\cos[(n-1)\theta/2] \cdot \sin(n\theta/2)}{\sin(\theta/2)}, \theta \neq 0, 2k\pi$$

$$S_n = \sum_{r=1}^{n} \sin r\theta = \frac{\sin[(n-1)\theta/2] \cdot \sin(n\theta/2)}{\sin(\theta/2)}, \theta \neq 0, 2k\pi.$$

Thus S_n is a bounded function, *i.e.*, the series $\Sigma \cos n\theta$ or $\Sigma \sin n\theta$ is such that its partial sums are bounded, when θ is neither zero nor a multiple of 2π. By Dirichlet's test, the series converges for $\alpha > 0$. For $\alpha \leq 0$ since the nth term does not tend to zero, both the series diverge.

For $\theta = 0$ or $2k\pi$, the series $\Sigma[(\cos n\theta)/n^\alpha]$ reduces to $\Sigma(1/n^\alpha)$ which is convergent for $\alpha > 1$, the series $\Sigma[(\sin n\theta)/n^\alpha]$ reduces to a series of zeros which is evidently convergent for all values of α.]

3. Show that the series

$$\sum_{n=1}^{\infty}\left(1+\frac{1}{2}+\frac{1}{3}+..., +\frac{1}{n}\right)\frac{\sin n\theta}{n}$$

converges, absolutely for $\theta = k\pi$, k is any integer, and conditionally for all other real values of θ.

4. Examine the convergence of the series:

(i) $\Sigma\left\{\frac{1}{(n+1)^2}+\frac{1}{(n+2)^2}+...+\frac{1}{(2n)^2}\right\}\cos n\theta$,

(ii) $\Sigma\left\{\frac{1}{(n+1)^2}+\frac{1}{(n+2)^2}+...+\frac{1}{(2n)^2}\right\}\sin^2 n\theta$.

5 Functions of a Single Variable (I)

LIMIT AND CONTINUITY

In the chapter on sequences, we considered functions whose domain was the set **N** of natural numbers. We shall now consider real valued functions with domain as any interval, open or closed.

1. LIMITS

Let f be a function defined for all points in a neighbourhood N of a point c except possibly at the point c itself.

Definition 1. The function f is said to tend to a limit l as x tends to (or approaches) c if for each $\varepsilon > 0$, there exists a $\delta > 0$ such that

$$0 < |x - c| < \delta \Rightarrow |f(x) - l| < \varepsilon$$

or

$$|f(x) - l| < \varepsilon, \text{ when } 0 < |x - c| < \delta$$

or

$$f(x) \in \left]l - \varepsilon, l + \varepsilon\right[, \forall x \in \left]c - \delta, c + \delta\right[\text{ except possibly } c$$

In symbols, we then write

$$\lim_{x \to c} f(x) = l$$

Definition 2. The function f is said to tend to $+\infty$ as x tends to c (or in symbols, $\lim_{x \to c} f(x) = +\infty$) if for each $G > 0$ (however large) there exists a $\delta > 0$ such that

$$f(x) > G, \text{ when } |x - c| < \delta$$

The function f is said to tend to $-\infty$ as x tends to c (or in symbols, $\lim_{x \to c} f(x) = -\infty$), if for each $G > 0$ (however large) there exists a $\delta > 0$ such that

$$f(x) < -G, \text{ when } |x - c| < \delta$$

Definition 3. The function f is said to tend to a limit l as x tends to ∞ (or in symbols, $\lim_{x \to \infty} f(x) = l$) if for each $\varepsilon > 0$, there exists a $k > 0$, such that

$$|f(x) - l| < \varepsilon, \text{ when } x > k$$

Definition 4. The function f is said to tend to $+\infty$ as x tends to ∞ (or in symbols, $\lim_{x\to\infty} f(x) = \infty$) if for each $G > 0$ (however large) there exists a $k > 0$, such that

$$f(x) > G, \text{ when } x > k$$

1.1 Left Hand and Right Hand Limits

While defining the limit of a function f as x tends to c, we consider values of $f(x)$ when x is very close to c. The values of x may be greater or less than c. If we restrict x to values less than c, then we say that x tends to c from below or from the left and write it symbolically as $x \to c-0$ or simply $x \to c-$. The limit of f with this restriction on x, is called the *left hand limit*. Similarly, if x takes only values greater than c, then x is said to tend to c from above or from the right, and is denoted symbolically as $x \to c+0$ or $x \to c+$. The limit of f is then called the *right hand limit*.

Definition 5. A function f is said to tend to a limit l as x tends to c from the left if for each $\varepsilon > 0$, there exists $\delta > 0$ such that

$$\left|f(x) - l\right| < \varepsilon, \text{ when } c-\delta < x < c$$

In symbols, we then write

$$\lim_{x\to c-0} f(x) = l \text{ or } f(c-0) = l \text{ or } f(c-) = l$$

Definition 6. A function f is said to tend to a limit l as x tends to c from the right if for each $\varepsilon > 0$, there exists $\delta > 0$ such that

$$\left|f(x) - l\right| < \varepsilon, \text{ when } c < x < c+\delta$$

In symbols, we then write

$$\lim_{x\to c+0} f(x) = l \text{ or } f(c+0) = l \text{ or } f(c+) = l$$

It may be noted that $\lim_{x\to c} f(x)$ exists if and only if both the limits, the left hand and the right hand, exist and are equal.

One-sided infinite limit may also be defined in the same way as above.

Example 1. Find the right hand and the left hand limits of a function defined as follows:

$$f(x) = \begin{cases} \dfrac{|x-4|}{x-4}, & x \neq 4, \\ 0, & x = 4. \end{cases}$$

- Now, when $x > 4$, $|x-4| = x-4$.

$$\therefore \quad \lim_{x\to 4+0} f(x) = \lim_{x\to 4+0} \frac{|x-4|}{x-4} = \lim_{x\to 4+0} \frac{x-4}{x-4} = \lim_{x\to 4+0} 1 = 1$$

Again when $x < 4$, $|x-4| = -(x-4)$.

$$\therefore \quad \lim_{x\to 4-0} f(x) = \lim_{x\to 4-0} \frac{-(x-4)}{x-4} = \lim_{x\to 4-0} (-1) = -1$$

so that

$$\lim_{x \to 4+0} f(x) \neq \lim_{x \to 4-0} f(x)$$

Hence $\lim\limits_{x \to 4+} f(x)$ does not exist.

Example 2. Evaluate $\lim\limits_{x \to 0+} \dfrac{1}{1 + e^{-1/x}}$.

■ [As $x \to 0+$, we feel that $1/x$ increases indefinitely, $e^{1/x}$ increases indefinitely. $e^{-1/x}$ tends to 0, $1 + e^{-1/x}$ tends to 1; thus the required limit may be 1.]

We have to show that for a given $\varepsilon > 0$, $\exists$ a $\delta > 0$ such that

$$\left| \frac{1}{1 + e^{-1/x}} - 1 \right| < \varepsilon, \text{ when } 0 < x < \delta$$

Now

$$\left| \frac{1}{1 + e^{-1/x}} - 1 \right| = \left| \frac{-e^{-1/x}}{1 + e^{-1/x}} \right| = \frac{1}{e^{1/x} + 1} < \varepsilon,$$

when $e^{1/x} + 1 > \dfrac{1}{\varepsilon}$ or $\dfrac{1}{x} > \log\left(\dfrac{1}{\varepsilon} - 1\right)$

$$\Rightarrow \qquad 0 < x < \frac{1}{\log(1/\varepsilon - 1)}, \text{ for } 0 < \varepsilon < 1$$

Thus choosing $\delta = \dfrac{1}{\log(1/\varepsilon - 1)}$, we see that if $0 < \varepsilon < 1$,

$$\left| \frac{1}{1 + e^{-1/x}} - 1 \right| < \varepsilon, \text{ when } 0 < x < \delta$$

Again when $\varepsilon \geq 1$,

$$\left| \frac{1}{1 + e^{-1/x}} - 1 \right| < \varepsilon \Rightarrow e^{1/x} > \frac{1}{\varepsilon} - 1$$

which is true for all values of x, so that any $\delta > 0$ would work.

Thus for any $\varepsilon > 0$ we are able to find a $\delta > 0$ such that

$$\left| \frac{1}{1 + e^{-1/x}} - 1 \right| < \varepsilon, \text{ when } 0 < x < \delta$$

$$\therefore \qquad \lim_{x\to 0+} \frac{1}{1+e^{-1/x}} = 0.$$

Example 3. Prove that $\lim_{x\to 0} x\sin\frac{1}{x} = 0$.

■ Now,

$$\left|x\sin\frac{1}{x}\right| = |x|\cdot\left|\sin\frac{1}{x}\right| \le |x|$$

Thus choosing a $\delta = \varepsilon$, we see that

$$\left|x\sin\frac{1}{x}\right| < \varepsilon, \text{ when } 0 < |x| < \delta$$

$$\Rightarrow \qquad \lim_{x\to 0} x\sin\frac{1}{x} = 0$$

Example 4. If $\lim_{x\to a} f(x)$ exists, prove that it must be unique.

■ Let, if possible, $f(x)$ tend to limits l_1 and l_2.

Hence for any $\varepsilon > 0$, it is possible to choose a $\delta > 0$ such that

$$|f(x)-l_1| < \frac{1}{2}\varepsilon, \text{ when } 0 < |x-a| < \delta$$

$$|f(x)-l_2| < \frac{1}{2}\varepsilon, \text{ when } 0 < |x-a| < \delta$$

Now

$$\begin{aligned} |l_1 - l_2| &= |l_1 - f(x) + f(x) - l_2| \\ &\le |l_1 - f(x)| + |f(x) - l_2| < \varepsilon, \end{aligned}$$

when $0 < |x-a| < \delta$

i.e., $|l_1 - l_2|$ is less than any positive number ε (however small) and so must be equal to zero. Thus $l_1 = l_2$.

Example 5. Show that $\lim_{x\to 3} \frac{1}{(x-3)^4} = \infty$.

■ Let G be any positive number, however large.

Now $\left|\frac{1}{(x-3)^4}\right| > G$

or

$$\frac{1}{(x-3)^4} > G, \text{ when } (x-3)^4 < \frac{1}{G} \text{ or when } 0 < |x-3| < \frac{1}{G^{1/4}}$$

Choosing $\delta = \frac{1}{G^{1/4}}$, we get the required result.

Example 6. Prove that $\lim_{x\to 0} \log|x| = -\infty$.

■ Given $G > 0$, choose $\delta = e^{-G}$. Now if $0 < |x-0| < \delta$ we have $|x| < e^{-G}$, and so $\log|x| < -G$, consequently $\lim_{x\to 0} \log|x| = -\infty$.

Example 7. Show that $\lim_{x\to 1} 2^{1/(x-1)}$ does not exist.

■ We first consider the left hand limit. Let $\varepsilon > 0$ be given. Choose a positive integer m such that, $1/2^m < \varepsilon$.

Take $\delta = \frac{1}{m}$ and let x satisfy $1-\delta < x < 1$. Now $-\delta < (x-1) < 0$, and so $\frac{1}{x-1} < -\frac{1}{\delta} < 0$.

Thus $\left|2^{1/(x-1)} - 0\right| = 2^{1/(x-1)} < 2^{-1/\delta} < 2^{-m} < \varepsilon$

and hence $\lim_{x\to 1-} 2^{1/(x-1)} = 0$.

Next, consider x to be on the right of 1.

Let $\delta > 0$ be arbitrary and choose a positive integer m_0 such that $\frac{1}{m_0} < \delta$. Then if $n \geq m_0$, $1 + \frac{1}{n} \in]1, 1+\delta[$ and $2^{\frac{1}{1+\frac{1}{n}-1}} = 2^n$, which is unbounded. Therefore $\lim_{x\to 1+} 2^{1/(x-1)}$ does not exist.

1.2 Theorems on Limits

Let f and g be two real functions with domain D. We define four new functions $f \pm g$, fg, f/g on domain D by setting

$$(f+g)\,x = f(x) + g(x)$$
$$(f-g)\,x = f(x) - g(x)$$
$$(f\cdot g)\,x = f(x)\cdot g(x)$$
$$(f/g)\,x = f(x)/g(x), \text{ if } g(x) \neq 0 \textit{ for any } x \in D$$

We shall now prove a theorem concerning the limits of these functions.

Theorem 1. *If f and g are two functions defined on some neighbourhood of c such that*

$$\lim_{x\to c} f(x) = l,\ \lim_{x\to c} g(x) = m$$

then

(*i*) $\lim\limits_{x\to c}(f + g)x = \lim\limits_{x\to c} f(x) + \lim\limits_{x\to c} g(x) = l + m\cdot$

(*ii*) $\lim\limits_{x\to c}(f - g)x = \lim\limits_{x\to c} f(x) - \lim\limits_{x\to c} g(x) = l - m.$

(*iii*) $\lim\limits_{x\to c}(f\cdot g)x = \lim\limits_{x\to c} f(x)\cdot \lim\limits_{x\to c} g(x) = lm.$

(*iv*) $\lim\limits_{x\to c}(f/g)x = \lim\limits_{x\to c} f(x)/\lim\limits_{x\to c} g(x) = l/m,$ *if* $m \neq 0.$

(*i*) Since $\lim\limits_{x\to c} f(x) = l,\ \lim\limits_{x\to c} g(x) = m,$ therefore for any $\varepsilon > 0, \exists$ positive numbers δ_1, δ_2 such that

$$|f(x) - l| < \frac{1}{2}\varepsilon, \text{ when } 0 < |x - c| < \delta_1$$

$$|g(x) - m| < \frac{1}{2}\varepsilon, \text{ when } 0 < |x - c| < \delta_2.$$

If $\delta = \min(\delta_1, \delta_2)$, then for $0 < |x - c| < \delta$,

$$\left|f(x) - l < \frac{1}{2}\right|\varepsilon,\ |g(x) - m| < \frac{1}{2}\varepsilon$$

and

$$\begin{aligned} |(f + g)x - (l + m)| &= |f(x) - l + g(x) - m| \\ &\le |f(x) - l| + |g(x) - m| \\ &< \varepsilon \end{aligned}$$

$\Rightarrow$ $|(f + g)x - (l + m)| < \varepsilon$, when $0 < |x - c| < \delta$

$\Rightarrow$ $\lim\limits_{x\to c}(f + g)x = l + m$

(*ii*) Proof is similar to part (*i*).

(*iii*) Let $\varepsilon > 0$ be given.

Now

$$\begin{aligned} |(f\cdot g)x - lm| &= |f(x)(g(x) - m) + m(f(x) - l)| \\ &\le |f(x)|\cdot|g(x) - m| + |m|\cdot|f(x) - l| \end{aligned} \quad \text{...(1)}$$

Since $\lim\limits_{x\to c} f(x) = l$, therefore for $\varepsilon = 1, \exists$ a $\delta_1 > 0$, such that

$$|f(x) - l| < 1, \text{ when } 0 < |x - c| < \delta_1.$$

Now

$$\left|f(x)\right| = \left|f(x) - l + l\right| \le \left|f(x) - l\right| + \left|l\right|$$

$$< 1 + |l|, \text{ when } 0 < |x - c| < \delta_1 \qquad \text{...(2)}$$

Again, since $\lim_{x \to c} g(x) = m$, therefore $\exists$ a $\delta_2 > 0$ such that

$$\left|g(x) - m\right| < \frac{\frac{1}{2}\varepsilon}{1 + |l|}, \text{ when } 0 < |x - c| < \delta_2. \qquad \text{...(3)}$$

Also, as $\lim_{x \to c} f(x) = l$, therefore $\exists$ a $\delta_3 > 0$ such that

$$\left|f(x) - l\right| < \frac{\frac{1}{2}\varepsilon}{1 + |m|}, \text{ when } 0 < |x - c| < \delta_3. \qquad \text{...(4)}$$

Let $\delta = \min(\delta_1, \delta_2, \delta_3)$.

Then, from (1), (2), (3) and (4), we have

$$\left|(f \cdot g)x - lm\right| < \frac{(1 + |l|)\frac{1}{2}\varepsilon}{1 + |l|} + \frac{|m|\varepsilon/2}{1 + |m|} < \varepsilon,$$

when $0 < |x - c| < \delta$.

Hence

$$\lim_{x \to c} (fg)(x) = lm$$

(*iv*) *Lemma.* Show that if $\lim_{x \to c} g(x) = m > 0$, then $\exists$ a $\delta_1 > 0$ such that

$$\left|g(x)\right| > \frac{1}{2}|m|, \text{ when } 0 < |x - c| < \delta_1$$

Since $\lim_{x \to c} g(x) = m$, therefore for $\varepsilon = \frac{1}{2}|m| > 0$, $\exists$ a $\delta_1 > 0$ such that

$$\left|g(x) - m\right| < \frac{1}{2}|m|, \text{ when } 0 < |x - c| < \delta_1$$

Also

$$|m| = \left|m - g(x) + g(x)\right| \le \left|m - g(x)\right| + \left|g(x)\right|$$

$$< \frac{1}{2}|m| + \left|g(x)\right|, \text{ when } 0 < |x - c| < \delta_1$$

i.e., $\quad \left|g(x)\right| > \frac{1}{2}|m|, \text{ when } 0 < |x - c| < \delta_1$

$\Rightarrow \quad \exists$ a deleted neighbourhood of c on which $g(x)$ does not vanish.

Let us now attend to the main theorem.

$$\left|(f/g)x - l/m\right| = \frac{mf(x) - lg(x)}{mg(x)}$$

$$\leq \frac{\left|f(x) - l\right|}{\left|g(x)\right|} + \frac{|l| \cdot \left|g(x) - m\right|}{|m| \cdot \left|g(x)\right|}$$

$$< \frac{2\left|f(x) - l\right|}{|m|} + \frac{2|l| \cdot \left|g(x) - m\right|}{|m|^2} \qquad ...(1)$$

where $0 < |x - c| < \delta_1$

Let $\varepsilon > 0$ be given.

Since $\lim_{x \to c} f(x) = l$, $\lim_{x \to c} g(x) = m$, therefore $\exists$ positive numbers δ_2 and δ_3 such that

$$\left|f(x) - l\right| < \frac{1}{4}\varepsilon|m|, \text{ when } 0 < |x - c| < \delta_2 \qquad ...(2)$$

and

$$\left|g(x) - m\right| < \frac{1}{4}\varepsilon\frac{|m|^2}{|l| + 1}, \text{ when } 0 < |x - c| < \delta_3 \qquad ...(3)$$

Let $\delta = \min(\delta_1, \delta_2, \delta_3)$.

Thus from (1), (2) and (3), we have

$$\left|(f/g)x - l/m\right| < \varepsilon, \text{ when } 0 < |x - c| < \delta$$

Hence $\lim_{x \to c} (f/g)x = l/m$, provided $m \neq 0$.

Example 8. Evaluate:

(*i*) $\lim_{x \to -1} \frac{(x + 2)(3x - 1)}{x^2 + 3x - 2}$

(*ii*) $\lim_{x \to 0} \frac{\sqrt{4 + x} - 2}{x}$

(*iii*) $\lim_{x \to 0+} \frac{\sin x}{\sqrt{x}}$.

■ (*i*) $\lim_{x \to -1} \frac{(x + 2)(3x - 1)}{x^2 + 3x - 2} = \frac{\lim_{x \to -1} (x + 2) \cdot \lim_{x \to -1} (3x - 1)}{\lim_{x \to -1} \left(x^2 + 3x - 2\right)}$

$$= \frac{1 \cdot (-4)}{-4} = 1.$$

(*ii*) $$\lim_{x \to 0} \frac{\sqrt{4+x}-2}{x} = \lim_{x \to 0} \frac{\sqrt{4+x}-2}{x} \cdot \frac{\sqrt{4+x}+2}{\sqrt{4+x}+2}$$

$$= \lim_{x \to 0} \frac{1}{\sqrt{4+x}+2} = \frac{1}{4}.$$

(*iii*) $$\lim_{x \to 0+} \frac{\sin x}{\sqrt{x}} = \left(\lim_{x \to 0+} \frac{\sin x}{x}\right) \cdot \left(\lim_{x \to 0+} \sqrt{x}\right) = 1 \cdot 0 = 0.$$

Example 9. Evaluate $\lim\limits_{x \to 1} \dfrac{x^2-1}{x-1}$.

■ Let us evaluate the left hand and the right hand limits.

When $x \to 1-0$, put $x = 1-h$, $h > 0$.

Hence $h \to 0+$ as $x \to 1-0$, so that

$$\lim_{x \to 1-0} \frac{x^2-1}{x-1} = \lim_{h \to 0+} \frac{(1-h)^2-1}{-h} = \lim_{h \to 0+} \frac{-h(2-h)}{-h}$$

$$= \lim_{h \to 0+} (2-h) = 2$$

Again when $x \to 1+0$, put $x = 1+h$, $h > 0$.

$$\therefore \quad \lim_{x \to 1+0} \frac{x^2-1}{x-1} = \lim_{h \to 0+} \frac{(1+h)^2-1}{h} = \lim_{h \to 0+} (2+h) = 2.$$

So that both, the left hand and the right hand, limits exist and are equal. Hence limit of the given function exists and equals 2.

Note: Since $x \neq 1$, division by $(x-1)$ is permissible.

$$\therefore \quad \lim_{x \to 1} \frac{x^2-1}{x-1} = \lim_{x \to 1} (x+1) = 2.$$

Example 10. Evaluate $\lim\limits_{x \to 0} \dfrac{e^{1/x}}{e^{1/x}+1}$.

■ Now when $x \to 0+$, $1/x \to \infty$, $e^{-1/x} \to 0$ and when $x \to 0-$, $1/x \to -\infty$, $e^{1/x} \to 0$.

$$\therefore \quad \lim_{x \to 0+} \frac{e^{1/x}}{e^{1/x}+1} = \lim_{x \to 0+} \frac{1}{1+e^{-1/x}} = 1$$

and

$$\lim_{x\to 0-} \frac{e^{1/x}}{e^{1/x}+1} = \frac{0}{1} = 0$$

so that the left hand limit is not equal to the right hand limit.

Hence $\lim_{x\to 0} \frac{e^{1/x}}{e^{1/x}+1}$ does not exist.

Example 11. Find $\lim_{x\to 0} e^x \operatorname{sgn}(x + [x])$, where the signum function is defined as

$$\operatorname{sgn}(x) = \begin{cases} 1, \text{ if } x > 0 \\ 1, \text{ if } x = 0 \\ -1, \text{ if } x < 0 \end{cases}$$

and $[x]$ means the greatest integer $\le x$.

■ L.H.L. $= \lim_{h\to 0+} e^{0-h} \operatorname{sgn}[0-h+(0-h)] = \lim_{h\to 0+} \left(-e^{-h}\right) = -1$

R.H.L. $= \lim_{h\to 0+} e^{0+h} \operatorname{sgn}[0+h+(0+h)] = \lim_{h\to 0+} e^{h} = 1$

$\therefore$ $\lim_{x\to 0} e^x \operatorname{sgn}(x + [x])$ does not exist.

EXERCISE

Evaluate:

1. $\lim_{x\to 2} \frac{x^2-4}{x-2}$

2. $\lim_{x\to 0-} \frac{|x|}{x}$ and $\lim_{x\to 0+} \frac{|x|}{x}$

3. $\lim_{x\to 0} \frac{3x+|x|}{7x-5|x|}$

4. $\lim_{x\to 1} \frac{1}{(x-1)}\left(\frac{1}{x+3} - \frac{2}{3x+5}\right)$

5. $\lim_{x\to 0} \frac{1-2\cos x+\cos 2x}{x^2}$

6. Show that $\lim_{x\to 0} \frac{xe^{1/x}}{1+e^{1/x}} = 0$.

7. Show that $\lim_{x \to 0} \frac{e^{1/x} - 1}{e^{1/x} + 1}$ does not exist.

8. Show that $\lim_{x \to 0} \frac{e^{1/x} - e^{-1/x}}{e^{1/x} + e^{-1/x}}$ does not exist.

9. If $\lim_{x \to c} f(x) = l$, then show that $\lim_{x \to c} |f(x)| = |l|$.

[*Hint.* $\big||f(x)| - |l|\big| \le |f(x) - l|, \forall x$

10. If $\lim_{x \to c} g(x) = m$ and $m \neq 0$, then show that $\lim_{x \to c} \frac{1}{g(x)} = \frac{1}{m}$.

ANSWERS

1. 4;
2. –1, 1;
3. Does not exist;
4. 1/32;
5. –1.

1.3 Limit of a Function (*Sequential approach*)

We have defined the limit of a function *f*, defined in a neighbourhood (taken as an interval) of a point *c*, to be *l* as *x* approaches *c*, if for any $\varepsilon > 0 \; \exists \; \delta > 0$ such that

$$|f(x) - l| < \varepsilon, \text{ when } 0 < |x - c| < \delta \qquad \text{...(1)}$$

We now give an alternative definition in terms of limits of sequences.

Definition. A number *l* is called the limit of a function *f* as *x* tends to *c* if the limit of the sequence $\{f(x_n)\}$ exists and is equal to *l* for any sequence $\{x_n\}$, $x_n \neq c$ for any *n*, convergent to *c*. Thus

$$\lim_{\substack{x_n \to c \\ x_n \neq c}} f(x) = l \qquad \text{...(2)}$$

Here, as in all similar cases, it is tactly implied that x_n *tending to c runs through a set of values for which f is defined.*

We now show that the two definitions are equivalent.

Let *f* have a limit *l* as $x \to c$ in the sense of the first definition, so that for any $\varepsilon > 0$, there exists $\delta > 0$, such that

$$|f(x) - l| < \varepsilon, \text{ for } 0 < |x - c| < \delta$$

Again, let there be a sequence $\{x_n\}$, $x_n \neq c$ for any *n*, converging to *c*, so that there exists a natural number *m* such that

$$|x_n - c| < \delta, \text{ for } n \ge m$$

Hence,

$$|f(x_n) - l| < \varepsilon, \text{ for } n \ge m,$$

i.e., the sequence $\{f(x_n)\}$ converges to l. Also, since this property holds for every sequence $\{x_n\}$, $x_n \neq c$, tending to c, we have proved that the limit exists in the sense of the second definition too.

Conversely, let a function f have a limit l and x tends to c in the sense of the second definition and suppose that it has no limit in the sense of the first definition. Then there exists at least one value of ε, say $\in_0$ for which there is no δ of the first definition, this means that for any δ there is a value $x = x^{(\delta)}$ belonging to the set satisfying $0 < |x - c| < \delta$ such that $\left|f\left(x^{(\delta)}\right) - l\right| \geq \varepsilon_0$.

Let δ take up successively the values $1, \frac{1}{2}, \frac{1}{3}, \ldots$ For each of them there is value x_k such that

$$|x_k - c| < \frac{1}{k}\,(x_k \neq c)$$

and

$$|f(x_k) - l| \geq \in_0 (k = 1, 2, 3, \ldots)$$

These relations show that $x_k\,(x_k \neq c)$ tends to c while $f(x_k)$ does not tend to l, which is a contradiction. Thus our supposition that the limit does not exist in the sense of the first definition is disproved.

We have thus proved the equivalence of the two definitions.

We now prove a very useful theorem due to Cauchy.

Cauchy's Criterion for Finite Limits

Theorem 2. *A function f tends to a finite limit as x tends to c if and only if for every* $\varepsilon > 0$, *there exists a neighbourhood N of c such that*

$$|f(x') - f(x'')| < \varepsilon, \text{ for all } x', x'' \in N;\ x', x'' \neq c$$

Necessary. Let $\lim_{x \to c} f(x) = l$, a finite number.

Therefore, for any $\varepsilon > 0$, there exists a deleted neighbourhood $N(c)$ such that $|f(x) - l| < \frac{1}{2}\varepsilon$, for $x \in N(c)$. Let $x', x'' \in N(c)$ so that

$$|f(x') - f(x'')| \leq |f(x') - l| + |l - f(x'')| < \frac{1}{2}\varepsilon + \frac{1}{2}\varepsilon = \varepsilon$$

Sufficient. Let for any $\varepsilon > 0$, there exist a deleted neighbourhood $N(c)$ of c such that $|f(x') - f(x'')| < \varepsilon$, for all $x', x'' \in N(c)$.

Let $\{x_n\}$, $x_n \neq c$ for any n, be an arbitrary sequence tending to c such that there exists a natural number m_0 such that $x_n, x_m \in N(c)$ for $n, m \geq m_0$.

Then

$$|f(x_n) - f(x_m)| < \varepsilon, \text{ for } n, m \geq m_0$$

Consequently, by Cauchy's general principle, sequence $\{f(x_n)\}$ tends to a limit.

We have thus proved for any sequence of numbers (x_n), $x_n \neq c$, converging to c, $\lim f(x_n)$ exists. Now, we prove that all these limits, $\lim f(x_n)$, corresponding to all possible different sequences tending to c, are equal to each other.

Let, if possible, $\{x_n\}$ and $\{x_n'\}$, $x_n \neq c$, $x_n' \neq c$ be two sequences tending to c such that sequences $\{f(x_n)\}$ and $\{f(x_n')\}$ tend to l and l' respectively. Let us construct the sequence $\{x_1, x_1', x_2, x_2', \ldots\}$ which converges to c. Therefore, by what has been proved above, the sequence $\{f(x_1), f(x_1'), f(x_2), \ldots\}$ converges, which is only possible if $l = l'$.

Hence the theorem.

We may similarly prove that:

A function f tends to a finite limit as x tends to ∞ *if and only if for every* $\in > 0$ *there exists* $G > 0$ *such that*

$$\left| f(x') - f(x'') \right| < \varepsilon, \text{ for all } x', x'' > G$$

Example 12. Show that $\lim\limits_{x \to 0} \dfrac{1}{x} \sin \dfrac{1}{x}$ does not exist.

■ Let $f(x) = \dfrac{1}{x} \sin \dfrac{1}{x}$. The function f is defined for every real number $x \neq 0$. Now for each natural number n, let $x_n = \dfrac{2}{\pi(4n+1)}$, and so

$$f(x_n) = (4n+1)\pi/2 \sin(2n\pi + \pi/2) = (4n+1)\pi/2 \to \infty \text{ as } n \to \infty.$$

$$\therefore \quad \lim_{n \to \infty} f(x_n) = \infty, \text{ when } \{x_n\} = \left\{ \frac{2}{(4n+1)\pi} \right\} \text{ converges to zero.}$$

Again, by taking $x_n = 1/n\pi$, we see $f(x_n) = n\pi \cdot 0 = 0$ for every natural number n, and so $\lim\limits_{x \to 0} f(x) \neq \infty$.

Therefore, $\lim\limits_{x \to 0} f(x)$ does not exist.

Example 13. Find $\lim\limits_{x \to -\infty} \left(x^2 \, sgn\, (\cos x) \right)$.

■ Let $x = -2n\pi$, so when $x \to -\infty$, $n \to \infty$.

Now

$$x^2 \, sgn\, (\cos x) = (-2n\pi)^2 \, sgn \cos(-2n\pi) = 4n^2\pi^2$$

$$\therefore \quad \lim_{x \to -\infty} x^2 \operatorname{sgn}(\cos x) = \infty$$

Again, let $x = -(2n+1)\pi$

$\therefore \qquad x^2 \, sgn\,(\cos x) = \left(-(2n+1)\pi\right)^2 sgn \cos\left(-(2n+1)\pi\right) = -(2n+1)^2 \pi^2$

and so

$$\lim_{x \to -\infty} x^2 \, sgn\,(\cos x) = -\infty$$

Hence

$$\lim_{x \to -\infty} x^2 \, sgn\,(\cos x) \text{ does not exist.}$$

2. CONTINUOUS FUNCTIONS

Let f be a function defined on an interval $[a, b]$. We shall now consider the behaviour of f at points of $[a, b]$.

2.1 Continuity at a Point

Definition 1 (Continuity at an internal point). A function f is said to be continuous at a point c, $a < c < b$, if

$$\lim_{x \to c} f(x) = f(c)$$

In other words, the function is continuous at c, if for each $\varepsilon > 0$, $\exists\, \delta > 0$ such that

$$\left| f(x) - f(c) \right| < \varepsilon, \text{ when } \left| x - c \right| < \delta$$

Definition 2. A function f is said to be *continuous from the left* at c if

$$\lim_{x \to c-0} f(x) = f(c)$$

Also f is *continuous from the right at c* if

$$\lim_{x \to c+0} f(x) = f(c)$$

Clearly a function is continuous at c if and only if it is continuous from the left as well as from the right.

Definition 3 (Continuity at an end point). A function f defined on a closed interval $[a, b]$ is said to be continuous at the end point a if it is continuous from the right at a, i.e.

$$\lim_{x \to a+0} f(x) = f(a)$$

Also the function is continuous at the end point b of $[a, b]$ if

$$\lim_{x \to b-0} f(x) = f(b)$$

Thus a function f is continuous at a point c if

(*i*) *$\lim_{x \to c} f(x)$ exists, and*

(*ii*) *limit equals the value of the function at $x = c$.*

2.2 Continuity in an Interval

A function f is said to be continuous in an interval $[a, b]$ if it is continuous at every point of the interval.

2.3 Discontinuous Functions

A function is said to be *discontinuous* at a point c of its domain if it is not continuous there at c. The point c is then called a *point of discontinuity* of the function.

Types of discontinuities

(*i*) A function f is said to have *a removable discontinuity* at $x = c$ if $\lim_{x \to c} f(x)$ exists but is not equal to the value $f(c)$ (which may or may not exist) of the function. Such a discontinuity can be removed by assigning a suitable value to the function at $x = c$.

(*ii*) f is said to have a *discontinuity of the first kind* at $x = c$ if $\lim_{x \to c-0} f(x)$ and $\lim_{x \to c+0} f(x)$ both exist but are not equal.

(*iii*) f is said to have a *discontinuity of the first kind from the left* at $x = c$ if $\lim_{x \to c-0} f(x)$ exists but is not equal to $f(c)$.
Discontinuity of the first kind from the right is similarly defined.

(*iv*) f is said to have a *discontinuity of the second kind* at $x = c$ if neither $\lim_{x \to c-0} f(x)$ nor $\lim_{x \to c+0} f(x)$ exists.

(*v*) f is said to have a *discontinuity of the second kind from the left* at $x = c$ if $\lim_{x \to c-0} f(x)$ does not exist.
Discontinuity of the second kind from the right may be defined similarly.

2.4 Theorems on Continuity

Theorem 3. *If f, g be two functions continuous at a point c then the functions $f + g$, $f - g$, fg are also continuous at c and if $g(c) \neq 0$, then f/g is also continuous at c.*

The proof is left as an exercise.

Theorem 4. *A function f defined on an interval I is continuous at a point $c \in I$ iff for every sequence $\{c_n\}$ in I converging to c, we have*

$$\lim_{n \to \infty} f(c_n) = f(c)$$

First let us suppose that the function f is continuous at a point $c \in I$, and $\{c_n\}$ is a sequence in I such that $\lim_{n \to \infty} c_n = c$.

Since f is continuous at c, therefore for any given $\in > 0, \exists$ a $\delta > 0$ such that

$$\left|f(x)-f(c)\right|<\varepsilon, \text{ when } 0<\left|x-c\right|<\delta \qquad \ldots(1)$$

Again, since $\lim\limits_{n\to\infty} c_n = c$, therefore $\exists$ a positive integer m, such that

$$\left|c_n - c\right| < \delta, \forall\, n \geq m \qquad \ldots(2)$$

From (1), putting $x = c_n$, we have

$$\left|f(c_n)-f(c)\right|<\varepsilon, \text{ when } \left|c_n - c\right|<\delta$$

$$\Rightarrow \qquad \left|f(c_n)-f(c)\right|<\varepsilon, \quad \forall\, n\geq m \qquad \text{[using 2]}$$

$\Rightarrow$ The sequence $\{f(c_n)\}$ converges to $f(c)$

or
$$\lim_{n\to\infty} f(c_n) = f(c)$$

Let us now suppose that f is *not continuous* at c, we shall now show that though $\exists$ a sequence $\{c_n\}$ in I converging to c, yet the sequence $\{f(c_n)\}$ does not converge to $f(c)$.

Since f is not continuous at c, therefore, there exists an $\varepsilon > 0$ such that for every $\delta > 0, \exists$ an $x \in I$ such that

$$\left|f(x)-f(c)\right|\geq\varepsilon, \text{ when } \left|x-c\right|<\delta$$

$\therefore$ By taking $\delta = 1/n$, we find that for each positive integer n, there is a $c_n \in I$ such that

$$\left|f(c_n)-f(c)\right|\geq\varepsilon, \text{ when } \left|c_n - c\right|<\frac{1}{n}$$

Thus the sequence $\{f(c_n)\}$ does not converge to $f(c)$, while the sequence $\{c_n\}$ converges to c.

Notes:

1. If $\lim c_n = c \Rightarrow \lim f(c_n) \neq f(c)$, then f is not continuous at c.
2. Compare with § 1.3.

Example 14. Examine the following function for continuity at the origin.

$$f(x) = \begin{cases} \dfrac{xe^{1/x}}{1+e^{1/x}} & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases}$$

■ Now

$$\lim_{x\to 0-} f(x) = \lim_{x\to 0-} \frac{xe^{1/x}}{1+e^{1/x}} = 0$$

and

$$\lim_{x\to 0+} f(x) = \lim_{x\to 0+} \frac{x}{e^{-1/x}+1} = 0$$

Thus,

$$\lim_{x\to 0-} f(x) = \lim_{x\to 0+} f(x) = \lim_{x\to 0} f(x) = 0$$

Also

$$\lim_{x\to 0} f(x) = f(0)$$

Thus, the function is continuous at the origin.

Example 15. Show that the function defined as:

$$f(x) = \begin{cases} \dfrac{\sin 2x}{x} & \text{when } x \neq 0 \\ & \text{when } x = 0 \end{cases}$$

has removable discontinuity at the origin.

■ Now

$$\lim_{x\to 0} f(x) = \lim_{x\to 0} \frac{\sin 2x}{2x} \cdot 2 = 2$$

so that

$$\lim_{x\to 0} f(x) \neq f(0)$$

Hence the limit exists, but is not equal to the value of the function at the origin. Thus the function has a removable discontinuity at the origin.

Note: The discontinuity can be removed by redefining the function at the origin such as $f(0) = 2$.

Example 16. Show that the function defined by

$$f(x) = \begin{cases} x \sin 1/x, & \text{when } x \neq 0 \\ 0, & \text{when } x = 0 \end{cases}$$

is continuous at $x = 0$.

■ Now

$$\lim_{x\to 0} f(x) = \lim_{x\to 0} \left(x \sin\frac{1}{x} \right) = 0$$

so that

$$\lim_{x\to 0} f(x) = f(0)$$

Hence, f is continuous at $x = 0$.

Example 17. A function f is defined on **R** by

$$f(x) = \begin{cases} -x^2 & \text{if } x \leq 0 \\ 5x - 4 & \text{if } 0 < x \leq 1 \\ 4x^2 - 3x & \text{if } 1 < x < 2 \\ 3x + 4 & \text{if } x \geq 2 \end{cases}$$

Examine f for continuity at $x = 0, 1, 2$. Also discuss the kind of discontinuity, if any.

■ Now

$$\lim_{x\to 0-} f(x) = \lim_{x\to 0-} \left(-x^2\right) = 0$$

$$\lim_{x\to 0+} f(x) = \lim_{x\to 0+} (5x-4) = -4$$

so that

$$\lim_{x\to 0-} f(x) = f(0) \neq \lim_{x\to 0+} f(x)$$

Thus the function has a discontinuity of the first kind from the right at the origin.

Again

$$\lim_{x\to 1-} f(x) = \lim_{x\to 1-} (5x-4) = 1$$

$$\lim_{x\to 1+} f(x) = \lim_{x\to 1+} \left(4x^2-3x\right) = 1$$

so that

$$\lim_{x\to 1-} f(x) = \lim_{x\to 1+} f(x) = 1 = f(1)$$

$$\Rightarrow \qquad \lim_{x\to 1} f(x) = f(1)$$

Thus the function is continuous at $x = 1$.

Again

$$\lim_{x\to 2-} f(x) = \lim_{x\to 2-} \left(4x^2-3x\right) = 10$$

$$\lim_{x\to 2+} f(x) = \lim_{x\to 2+} (3x+4) = 10$$

Also

$$f(2) = 10 \Rightarrow \lim_{x\to 2} f(x) = f(2)$$

Thus, the function is continuous at $x = 2$.

Example 18. Is the function f, where $f(x) = \dfrac{x-|x|}{x}$ continuous?

■ For $x < 0$, $f(x) = \dfrac{x+x}{2} = 2$, continuous.

For $x > 0$, $f(x) = \dfrac{x-x}{x} = 0$, continuous.

The function is not defined at $x = 0$.

Thus $f(x)$ is continuous for all x except at zero.

Example 19. Discuss the kind of discontinuity, if any of the function is defined as follows:

$$f(x) = \begin{cases} \dfrac{x - |x|}{x} & \text{when } x \neq 0 \\ 2 & \text{when } x = 0 \end{cases}$$

■ The function is continuous at all points except possibly the origin.

Let us test at $x = 0$.

Now,

$$\lim_{x \to 0-} f(x) = \lim_{x \to 0-} \frac{x + x}{x} = 2$$

$$\lim_{x \to 0+} f(x) = \lim_{x \to 0+} \frac{x - x}{x} = 0$$

and
$$f(0) = 2$$

Thus the function has discontinuity of the first kind from the right at $x = 0$.

Example 20. If $[x]$ denotes the largest integer $\leq x$, then discuss the continuity at $x = 3$ for the function

$$f(x) = x - [x], \forall\, x \geq 0$$

■ Now

$$\lim_{x \to 3-} f(x) = \lim_{x \to 3-} \{x - [x]\} = 3 - 2 = 1$$

$$\lim_{x \to 3+} f(x) = \lim_{x \to 3+} \{x - [x]\} = 3 - 3 = 0$$

and

$$f(3) = 0$$

Thus the function has a discontinuity of the first kind from the left at $x = 3$.

Note: The function is continuous at all points except the integral values 1, 2, 3, ...

Example 21. Prove that the *Dirichlet's* function f defined on **R** by

$$f(x) = \begin{cases} 1, & \text{when } x \text{ is irrational} \\ -1, & \text{when } x \text{ is rational} \end{cases}$$

is discontinuous at every point.

■ First, let a be any *rational* number so that $f(a) = -1$.

Since in any interval there lie an infinite number of rational and irrational numbers, therefore, for each positive integer n, we can choose an irrational number a_n such that $|a_n - a| < \dfrac{1}{n}$.

Thus the sequence $\{a_n\}$ converges to a.

But $f(a_n) = 1$ for all n, and $f(a) = -1$, so that

$$\lim_{n \to \infty} f(a_n) \neq f(a)$$

Thus by Theorem 4, § 2.4, the function is discontinuous at any rational number a.

Hence, the function is discontinuous at all rational points.

Next, let b be any *irrational* number. For each positive integer n we can choose a rational number b_n such that $|b_n - b| < \frac{1}{n}$. Thus the sequence $\{b_n\}$ converges to b.

But $f(b_n) = -1$ for all n and $f(b) = 1$.

$$\therefore \qquad \lim_{n \to \infty} f(b_n) \neq f(b)$$

Hence, the function is discontinuous at all irrational points.

Example 22. Show that the function $f(x)$ defined on **R** by

$$f(x) = \begin{cases} x, \text{ when } x \text{ is irrational} \\ -x, \text{ when } x \text{ is rational} \end{cases}$$

is continuous only at $x = 0$.

■ First, let $a \neq 0$ be any *rational* number, so that $f(a) = -a$. Since in every interval there lie an infinite number of rational and irrational numbers, therefore, for each positive integer n, we can choose an irrational number a_n such that

$$|a_n - a| < \frac{1}{n}$$

Thus the sequence $\{a_n\}$ converges to a.

But

$$\lim_{n \to \infty} f(a_n) = \lim_{n \to \infty} a_n = a$$

Thus

$$\lim_{n \to \infty} f(a_n) \neq f(a), a \neq 0$$

so that by Theorem 4, § 2.4, the function is discontinuous at any rational number, other than zero.

In a similar way the function may be shown to be discontinuous at every *irrational* point.

It may be seen from above, that the function is continuous at $x = 0$ (*i.e.*, $a = 0$). However, it can be shown to be continuous at $x = 0$ as follows:

Let $\varepsilon > 0$ be given and let $\delta = \varepsilon$ (or any $\delta < \varepsilon$), then

$|x| < \delta \Rightarrow |f(x) - f(0)| = |-x| = |x| < \varepsilon$, when x is rational and

$|x| < \delta \Rightarrow |f(x) - f(0)| = |x| < \varepsilon$, when x is irrational

Thus

$$|x| < \delta \quad \Rightarrow \quad |f(x) - f(0)| < \varepsilon$$

or

$$\lim_{x \to 0} f(x) = f(0)$$

Hence, the function is continuous at $x = 0$.

EXERCISE

Investigate the continuity at the indicated point:

1. $f(x) = \begin{cases} \dfrac{\sin(x-c)}{x-c}, & \text{if } x \neq c \\ 0 & \text{, if } x = c \end{cases}$ at $x = c$

2. $f(x) = \begin{cases} \dfrac{\tan^{-1} x}{x}, & \text{if } x \neq 0 \\ 1 & \text{, if } x = 0 \end{cases}$ at $x = 0$

3. $f(x) = \begin{cases} \dfrac{x^3-8}{x^2-4}, & \text{if } x \neq 2 \\ 3 & \text{, if } x = 2 \end{cases}$ at $x = 2$

4. $f(x) = x - |x|$, at $x = 0$

5. $f(x) = \begin{cases} x^m \sin\dfrac{1}{x}, & \text{if } x \neq 0, m > 0 \\ 0 & \text{, if } x = 0 \end{cases}$ at $x = 0$

6. $f(x) = \begin{cases} (1+x)^{1/x}, & \text{if } x \neq 0 \\ 1 & \text{, if } x = 0 \end{cases}$ at $x = 0$

7. $f(x) = \begin{cases} \dfrac{e^{1/x^2}}{1-e^{1/x^2}}, & \text{when } x \neq 0 \\ 1 & \text{, when } x = 0 \end{cases}$ at $x = 0$

8. $f(x) = \begin{cases} \dfrac{e^{1/x}-e^{-1/x}}{e^{1/x}+e^{-1/x}}, & \text{when } x \neq 0 \\ 1 & \text{, when } x = 0 \end{cases}$ at $x = 0$

9. Examine the continuity at $x = 1$.

$$f(x) = \begin{cases} 2x, & \text{when } 0 \le x < 1 \\ 3, & \text{when } x = 1 \\ 4x, & \text{when } 1 < x \le 2 \end{cases}$$

10. Obtain the points of discontinuity of the function f, defined on $[0, 1]$ as follows:

$$f(0) = 0,\ f(x) = \frac{1}{2} - x,\ \text{if } 0 < x < \frac{1}{2},$$

$$f\left(\frac{1}{2}\right) = \frac{1}{2},\ f(x) = \frac{3}{2} - x,\ \text{if } \frac{1}{2} < x < 1$$

$$f(1) = 1$$

Also examine the kind of discontinuities.

CHAPTER 5

11. Show that the function f defined on **R** by

$$f(x) = \begin{cases} 1, & \text{if } x \text{ is rational} \\ 0, & \text{if } x \text{ is irrational} \end{cases}$$

is discontinuous at every point.

12. Show that the function f defined by

$$f(x) = \begin{cases} x, & \text{if } x \text{ is rational} \\ 0, & \text{if } x \text{ is irrational} \end{cases}$$

is continuous only at $x = 0$.

13. Show that the function f defined by

$$f(x) = \begin{cases} [x+1]\sin\dfrac{1}{x}, & x \in \left]-1,0\right[\cup\left]0,1\right[\\ 0, & \text{otherwise} \end{cases}$$

has discontinuity of the second kind at $x = 0$ and discontinuity of the first kind at $x = 1$.

14. Show that the function f defined by

$$f(x) = \begin{cases} (1+x)\operatorname{sgn} x + \operatorname{sgn}|x| - 1, & \text{if } x \text{ is rational} \\ \operatorname{sgn} x, & \text{if } x \text{ is irrational} \end{cases}$$

has discontinuity of the second kind at $x \neq 0$ and discontinuity of the first kind at $x = 0$.

ANSWERS

1. Removable discontinuity
2. Continuous
3. Continuous
4. Continuous
5. Continuous
6. Removable discontinuity
7. Removable discontinuity
8. Discontinuity of the first kind from the left
9. Discontinuity of the first kind
10. Discontinuity of the first kind from the right at 0, discontinuity of the first kind at $x = \frac{1}{2}$, discontinuity of the first kind from the left at $x = 1$.

3. FUNCTIONS CONTINUOUS ON CLOSED INTERVALS

We shall now study some properties of functions which are continuous on closed intervals. In fact, we shall show that a function which is continuous on a closed interval, is bounded, attains its bounds and assumes every value between the bounds.

Theorem 5. *If a function is continuous in a closed interval, then it is bounded therein.*

Let f be a function defined and continuous in a closed interval I.

We shall show that if the function f is not bounded, then it fails to be continuous at some point of the closed interval I.

Let, if possible, f is *not bounded above*, so that for each positive integer n $\exists$ a point $x_n \in I$ such that $f(x_n) > n$.

Now $\{x_n\}$, being a sequence in the closed interval I, is bounded and has at least one limit point, say, ξ.

A closed interval is a closed set and so $\xi \in I$.

Further, since ξ is a limit point of the sequence $\{x_n\}$, therefore, there exists a subsequence (x_{n_k}) of $\{x_n\}$ such that $x_{n_k} \to \xi$ as $k \to \infty$.

Also since $f(x_{n_k}) > n_k$, for all k, therefore the sequence $\{f(x_{n_k})\}$ diverges to ∞.

Thus $\exists$ a point ξ of I such that a sequence (x_{n_k}) in I converges to ξ, but

$$\lim_{k \to \infty} f\left(x_{n_k}\right) \neq f(\xi)$$

Thus f is not continuous at ξ, which is a contradiction and hence the function is bounded above.

By considering a function 'f', it can be shown in a similar way that the function f is also bounded below.

Hence, the function is bounded.

Theorem 6. *If a function f is continuous on a closed interval $[a, b]$, then it attains its bounds at least once in $[a, b]$.*

If f is a constant function, then evidently it attains its bounds at every point of the interval.

Let f be a function which is not a constant.

Since f is continuous on the closed interval $[a, b]$, therefore, it is bounded. Let m and M be the infimum and supremum of f. It is to be shown that $\exists$ point α, β of $[a, b]$ such that

$$f(\alpha) = m,\ f(\beta) = M$$

Let us consider the case of the *supremum*.

Suppose f does not attain the supremum M so that the function does not take the value M for any point $x \in [a, b]$, *i.e.*,

$$f(x) \neq M, \text{ for any } x \in [a, b]$$

Now consider the function

$$g(x) = \frac{1}{M - f(x)},\ \forall\, x \in [a, b]$$

which is positive for all values of x in $[a, b]$.

Evidently the function g is continuous and so bounded in $[a, b]$. Let k (>0) be its supremum.

$$\therefore \qquad \frac{1}{M - f(x)} \leq k,\ \forall\, x \in [a, b]$$

$$\Rightarrow \qquad f(x) \leq M - \frac{1}{k},\ \forall\, x \in [a, b]$$

which contradicts the hypothesis that M is the supremum of f in $[a, b]$. Hence our supposition that f does not attain the value M leads to a contradiction and therefore f attains its supremum for at least one value in $[a, b]$.

It may similarly be shown that the function also attains its infimum m.

Hence, the function attains its bounds at least once in $[a, b]$.

Note: It may be observed from the two preceding theorems, that the function f, continuous on the closed interval $[a, b]$, has the least and the greatest values m and M, *i.e.*, the range set of f is bounded with m and M as its smallest and greatest elements. Thus the range set of f is a subset of $[m, M]$. We shall, in fact, show later that the range set of f is $[m, M]$ itself and that f takes up every value between m and M.

ILLUSTRATIONS

1. The function $f(x) = \dfrac{1}{1+|x|}$, for real x, is continuous and bounded and attains its supremum for $x = 0$ but does not attain the infimum.
2. The function $f(x) = -\dfrac{1}{1+|x|}$, (for all $x \in \mathbf{R}$, is continuous and bounded, attains its infimum but not the supremum.
3. The function $f(x) = x$, for all $x \in]0, 1[$ is continuous and bounded but attains neither the infimum nor the supremum.

Theorem 7. *If a function f is continuous at an interior point c of an interval $[a, b]$ and $f(c) \neq 0$, then $\exists\, a\, \delta > 0$ such that $f(x)$ has the same sign as $f(c)$, for every $x \in]c-\delta, c+\delta[$.*

Since the function f is continuous at an interior point c of $[a, b]$, therefore for any $\varepsilon > 0, \exists\, a\, \delta > 0$ such that

$$|f(x) - f(c)| < \varepsilon, \ \forall\, x \in]c-\delta, c+\delta[$$

or

$$f(c) - \varepsilon < f(x) < f(c) + \varepsilon, \ \forall\, x \in]c-\delta, c+\delta[$$

When $f(c) > 0$, taking ε to be less than $f(c)$ we find that

$$f(x) > 0, \ \forall\, x \in]c-\delta, c+\delta[$$

When $f(c) < 0$, taking ε to be less than $-f(c)$ we find that

$$f(x) < 0, \ \forall\, x \in]c-\delta, c+\delta[$$

Hence the theorem.

Corollary. If f is continuous at the end point b of $[a, b]$ and $f(b) \neq 0$, then there exists an interval $]b-\delta, b[$ such that $f(x)$ has the sign of $f(b)$ for all x in $]b-\delta, b]$.

A similar result holds for continuity at a.

Note: When c is an interior point of the interval, the theorem may be restated as:
If a function f is continuous at an interior point c of an interval $[a, b]$ and $f(c) \neq 0$, then $\exists$ a neighbourhood N of c wherein $f(x)$ has the same sign as $f(c)$, for all $x \in N$.

Theorem 8. *If a function f is continuous on a closed interval [a, b] and f (a) and f (b) are of opposite signs (f (a) · f (b) < 0), then there exists at least one point* $\alpha \in \left]a, b\right[$ *such that* $f(\alpha) = 0$.

Let us suppose that $f(a) > 0$ and $f(b) < 0$.

Let S consists of those points of $[a, b]$ for which $f(x)$ is positive, *i.e.*,

$$S = \{x : a \le x \le b \wedge f(x) > 0\}$$

Now

$$f(a) > 0 \quad \Rightarrow \quad a \in S \Rightarrow S \ne \phi$$

Also S is bounded above by b.

Hence by the order completeness property, S has the supremum, say α, where $a \le \alpha \le b$.

We shall now show that

(*i*) $\alpha \ne a$, $\alpha \ne b$, and

(*ii*) $f(\alpha) = 0$.

(*i*) First we show that $\alpha \ne a$

Since $f(a) > 0$, therefore $\exists$ a $\delta > 0$ such that

$$f(x) > 0, \ \forall\, x \in [a, a + \delta[$$

$\Rightarrow$ $[a, a + \delta[\subseteq S$

$\Rightarrow$ the supremum α of S is greater than or equal to $a + \delta$

$\Rightarrow$ $\alpha \ne a$

Now we shall show that $\alpha \ne b$.

Since $f(b) < 0$, therefore $\exists$ a $\delta > 0$ such that

$$f(x) < 0, \ \forall\, x \in \left]b - \delta, b\right]$$

$\Rightarrow$ $b - \delta$ is an upper bound of S

$\Rightarrow$ $\alpha \le b - \delta \Rightarrow \alpha \ne b$

(*ii*) We shall now show that $f(\alpha) \not> 0$ and $f(\alpha) \not< 0$.

If $f(\alpha) > 0$, then $\exists$ a $\delta > 0$ such that

$$f(x) > 0, \ \forall\, x \in \left]\alpha - \delta, \alpha + \delta\right[$$

$\Rightarrow$ $\left]\alpha - \delta, \alpha + \delta\right[\subseteq S$

Let us choose a positive $\delta_2 < \delta$ such that $\alpha + \delta_2 \in \left]\alpha - \delta, \alpha + \delta\right[$

A member $\alpha + \delta_2$ of S is greater than the supremum α of S, which is a contradiction.

$\therefore$ $\qquad f(\alpha) \not> 0$

Let now $f(\alpha) < 0$, so that $\exists$ a $\delta_1 > 0$ such that

$$f(x) < 0, \ \forall\, x \in \left]\alpha - \delta_1, \alpha + \delta_1\right[\qquad ...(1)$$

Again, since α is the supremum of S, therefore, there exists a member β of S, where $\alpha - \delta_1 < \beta \le \alpha$ such that

$$f(\beta) > 0$$

But from (1), $f(\beta) < 0$, which is a contradiction.

$$\therefore \qquad f(\alpha) \nless 0$$

Thus, it follows that $f(\alpha) = 0$.

Theorem 9. ***Intermediate value theorem.*** *If a function f is continuous on* $[a, b]$ *and* $f(a) \neq f(b)$, *then it assumes every value between* $f(a)$ *and* $f(b)$.

Let A be any number between $f(a)$ and $f(b)$. We shall show that there exists a number $c \in]a, b[$ such that $f(c) = A$.

Consider a function ϕ defined on $[a, b]$ such that

$$\phi(x) = f(x) - A$$

Clearly $\phi(x)$ is continuous on $[a, b]$.

Also

$$\phi(a) = f(a) - A, \text{ and } \phi(b) = f(b) - A$$

so that $\phi(a)$ and $\phi(b)$ are of opposite signs.

Thus the function ϕ is continuous on $[a, b]$ and $\phi(a)$ and $\phi(b)$ are of opposite signs; therefore, by the previous theorem, $\exists\ c \in]a, b[$ such that

$$\phi(c) = 0$$

$$\Rightarrow \qquad f(c) - A = 0 \Rightarrow f(c) = A$$

Corollary. A function f, which is continuous on a closed interval $[a, b]$, assumes every value between its bounds.

Since the function f is continuous on the closed interval $[a, b]$, therefore, it is bounded and attains its bounds on $[a, b]$, *i.e.*, $\exists$ two numbers α, β in $[a, b]$ such that

$$f(\alpha) = M \text{ and } f(\beta) = m,$$

where M and m are, respectively, the supremum and the infimum of f.

Since f is continuous on $[a, b]$, therefore, it is continuous on $[\beta, \alpha]$ or $[\alpha, \beta]$ as the case may be, and consequently assumes every value between $f(\alpha)$ and $f(\beta)$.

Thus the function assumes every value between its bounds.

We may sum up in other words:

The range of a continuous function whose domain is a closed interval is as well a closed interval.

Or, in still better words:

The image of a closed interval under a continuous function (mapping) is a closed interval.

Example 23. Show that the function defined on $[0, 1]$ as

$$f(x) = 2x + 1, \forall\ x \in]0, 1]$$
$$f(0) = 0$$

does not satisfy the conclusion of the intermediate value theorem.

- The function f is bounded, but is not continuous on $[0, 1]$, since f fails to be right-continuous at $x = 0$.

$f(1) = 3 = M$ and $f(0) = 0 = m$, but there is no $c \in]0, 1[$ with $f(c) = 1$. It is easy to see that none of the intermediate values $x \in]0, 3[$ are assumed by f.

Theorem 10. ***Fixed point theorem.*** *If f is continuous on $[a, b]$ and $f(x) \in [a, b]$, for every $x \in [a, b]$ then f has a fixed point, i.e., there exists a point, $c \in [a, b]$ such that $f(c) = c$.*

Suppose f is continuous on $[a, b]$ and $f(x) \in [a, b]$ for every $x \in [a, b]$. If $f(a) = a$ or $f(b) = b$, then the theorem is proved, hence we assume that $f(a) > a$ and $f(b) < b$.

Let $g(x) = f(x) - x, \ \forall \ x \in [a, b]$

Now $g(a) > 0$, $g(b) < 0$ and g is continuous on $[a, b]$. '0' is an intermediate value of g on $[a, b]$. Hence by intermediate value theorem, there exists a point $c \in]a, b[$ such that $g(c) = 0$. Then $f(c) = c$.

Definition. A function f defined on $[a, b]$ is said to satisfy the *intermediate-value property* on $[a, b]$ if for every $x_1, x_2 \in [a, b]$ with $x_1 < x_2$ and for every A between $f(x_1)$ and $f(x_2)$ there is a $c \in]x_1, x_2[$ with $f(c) = A$.

A function which satisfies the intermediate-value property on $[a, b]$ need not be continuous on $[a, b]$. For example, the function $f(x) = \sin 1/x$ with $f(0) = 0$ defined on $[-2/\pi, 2/\pi]$ satisfies the intermediate value property but is not continuous at $x = 0$.

Ex. 1 If f satisfies the intermediate-value property on $[a, b]$, then prove that f has no discontinuities (of the first kind and removable on $[a, b]$).

Ex. 2 Prove that if f is one-to-one and satisfies the intermediate-value property on $[a, b]$, then f is continuous on $[a, b]$.

(*Hint*: Monotone functions have no discontinuities of the second kind).

Ex. 3 If f is continuous on $[a, b]$ and $f(a) = f(b)$, then show that there exist $x, y \in]a, b[$ such that $f(x) = f(y)$.

4. UNIFORM CONTINUITY

Let f be a function defined on an interval I. Then by definition, the function is continuous at any point $c \in I$ if for any $\varepsilon > 0$, there exists a $\delta > 0$ such that

$$|f(x) - f(c)| < \varepsilon, \text{ when } |x - c| < \delta.$$

For continuity at any other point $d \in I$, for the same ε, a $\delta_1 > 0$ would exist (not necessarily equal to δ). There is in fact a δ corresponding to each point of I. The number δ in general depends on the selection of ε and the point c. However, if a δ could be found which depends only on ε and not on the selection of the point c, such a δ would work for the whole interval I on which f is continuous. In such a case, f is said to be *uniformly continuous* on I. Thus, the notion of uniform continuity is *global* in character in as much as we talk of uniform continuity only on an interval.

The notion of continuity is, however, *local* in character in as much as we can talk of continuity at a point.

It may seem to a beginner that the infimum of the set consisting of δ's corresponding to different points of I would work for the whole of I. But the infimum may be zero. In general, therefore, a δ

which may work for the entire interval may not exist, so that every continuous function may not be uniformly continuous.

Definition. A function f defined on an interval I is said to be *uniformly continuous* on I if to each $\varepsilon > 0$ there exists a $\delta > 0$ such that

$$|f(x_2) - f(x_1)| < \varepsilon, \text{ for arbitrary points } x_1, x_2 \text{ of } I \text{ for which } |x_1 - x_2| < \delta$$

4.1 We shall now prove two theorems on uniform continuity.

Theorem 11. *A function which is uniformly continuous on an interval is continuous on that interval.*

Let a function f be uniformly continuous on an interval I, so that for a given $\varepsilon > 0$, there corresponds a $\delta > 0$ such that

$$|f(x_1) - f(x_2)| < \varepsilon, \text{ where } x_1, x_2 \text{ are any two points of } I \text{ for which}$$

$$|x_1 - x_2| < \delta$$

Let $x \in I$, then on taking $x_1 = x$, we find that for $\varepsilon > 0$, $\exists\ \delta > 0$ such that

$$\left|f(x) - f(x_2)\right| < \varepsilon, \text{ when } \left|x - x_2\right| < \delta.$$

Hence the function is continuous at every point $x_2 \in I$, *i.e.*, the function f is continuous on I.

Theorem 12. *A function which is continuous on a closed interval is also uniformly continuous on that interval.*

Let a function f be continuous on a closed interval I. Let, if possible, f be not uniformly continuous on I. Then there exists an $\varepsilon > 0$ such that for any $\delta > 0$, there are numbers $x, y \in I$ for which

$$\left|f(x) - f(y)\right| \not< \varepsilon, \text{ when } \left|x - y\right| < \delta$$

In particular for each positive integer n, we can find real numbers x_n, y_n in I such that

$$\left|f(x_n) - f(y_n)\right| \not< \varepsilon, \text{ when } \left|x_n - y_n\right| < 1/n \qquad \text{...(1)}$$

Now $\{x_n\}$ and $\{y_n\}$ being sequences in the closed interval I, they are bounded and so each has at least one limit point, say ξ and η respectively.

As a closed interval is a closed set,

$$\therefore \qquad \xi \in I, \eta \in I$$

Further since ξ is a limit point of $\{x_n\}$, there exists a convergent subsequence $\{x_{n_k}\}$ of $\{x_n\}$ such that $x_{n_k} \to \xi$ as $k \to \infty$.

Similarly, there exists a convergent subsequence $\{y_{n_k}\}$ of $\{y_n\}$ such that $y_{n_k} \to \eta$ as $k \to \infty$.

Again from (1), we find that

$$\left|f(x_{n_k}) - f(y_{n_k})\right| \not< \varepsilon, \text{ when } \left|x_{n_k} - y_{n_k}\right| < 1/n_k \le 1/k \qquad \text{...(2)}$$

The second inequality shows that

$$\lim_{k \to \infty} x_{n_k} = \lim_{k \to \infty} y_{n_k}$$

$$\Rightarrow \qquad \xi = \eta$$

From the first inequality we find that in case the sequences $\{f(x_{n_k})\}$ and $\{f(y_{n_k})\}$ converge, the limits to which they converge are different.

We thus have two sequences $\{x_{n_k}\}$ and $\{y_{n_k}\}$ both of which converge to ξ but $\{f(x_{n_k})\}$ and $\{f(y_{n_k})\}$ do not converge to the same limit.

So by Theorem 4 § 2.4, f is not continuous at ξ, for, otherwise, the two sequences $\{f(x_{n_k})\}$ and $\{f(y_{n_k})\}$ would converge to the same point $f(\xi)$.

Thus we arrive at a contradiction and so the hypothesis that f is not uniformly continuous on I is false.

Hence, f is uniformly continuous on I.

Example 24. Show that the function $f(x) = 1/x$ is not uniformly continuous on $]0, 1]$.

■ Clearly the function is continuous on $]0, 1]$.

It will be uniformly continuous on the given interval if for a given $\varepsilon > 0, \exists$ a $\delta > 0$, independent of the choice of points x and c in $]0, 1]$, such that

$$\left|\frac{1}{x} - \frac{1}{c}\right| < \varepsilon, \text{ when } |x - c| < \delta$$

or

$$\left|\frac{c - x}{cx}\right| < \varepsilon, \text{ when } c - \delta < x < c + \delta \qquad \text{...(1)}$$

If we take $c = \delta$, then the interval $]c - \delta, c + \delta[$ becomes $]0, 2\delta[$. Also condition (1) must hold for any x in this interval.

But

$$\frac{\delta - x}{\delta x} \to \infty \text{ as } x \to 0,$$

i.e., if we choose x sufficiently close to zero, then condition (1) is violated.

Hence, $1/x$ is not uniformly continuous on $]0, 1]$.

Aliter.

If $\varepsilon = \dfrac{1}{2}$ and δ is any positive number, then for $n > \dfrac{1}{\delta}$,

$$\left|\frac{1}{n} - \frac{1}{n+1}\right| = \frac{1}{n(n+1)} < \frac{1}{n} < \delta$$

Therefore taking $x_1 = \dfrac{1}{n}$ and $x_2 = \dfrac{1}{n+1}$, as any two points of the interval $]0, 1]$, we have

$$|f(x_1) - f(x_2)| = |n - (n+1)| = 1 > \varepsilon, \text{ whenever } |x_1 - x_2| < \delta.$$

Hence, f is not uniformly continuous on $]0, 1]$.

Note: The function $f(x) = 1/x$ is uniformly continuous on $[a, \infty[$ where $a > 0$.

Example 25. Prove that $f(x) = \sin x^2$ is not uniformly continuous on $[0, \infty[$.

■ Let $\varepsilon = \frac{1}{2}$ and δ be any positive number such that for $n > \pi/\delta^2$

$$\left|\sqrt{\frac{n\pi}{2}} - \sqrt{\frac{(n+1)\pi}{2}}\right| < \delta$$

Therefore, taking $x_1 = \sqrt{\frac{n\pi}{2}}$ and $x_2 = \sqrt{\frac{(n+1)\pi}{2}}$, as any two points of the interval $[0, \infty[$

$$|f(x_1) - f(x_2)| = \left|\sin\frac{n\pi}{2} - \sin\frac{(n+1)\pi}{2}\right| = 1 > \varepsilon,$$

$$|x_1 - x_2| < \delta$$

Hence $f(x) = \sin x^2$ is not uniformly continuous on $[0, \infty[$.

Example 26. Prove that

$$\begin{aligned} f(x) &= \sin\frac{1}{x}, x \neq 0 \\ &= 0, \qquad x = 0 \end{aligned}$$

is not uniformly continuous on $[0, \infty[$.

■ Let $\varepsilon = \frac{1}{2}$ and $\delta > 0$ be such that $\frac{1}{n(2n\pi + \pi)} < \delta,\ \forall\, n \geq m$. Taking $x = \frac{2}{2m\pi}$ and $y = \frac{2}{(2m+1)\pi}$ be any two points of $[0, \infty[$, then

$$|x - y| = \left|\frac{2}{2m\pi} - \frac{2}{(2m+1)\pi}\right| = \frac{1}{m(2m+1)\pi} < \delta$$

does not imply

$$|f(x) - f(y)| = \left|\sin m\pi - \sin\left(m\pi + \frac{\pi}{2}\right)\right| = 1 < \varepsilon$$

Hence, $\sin\frac{1}{x}$ is not uniformly continuous on $]0, \infty[$.

Example 27. Show that the function $f(x) = x^2$ is uniformly continuous on $[-1, 1]$.

■ Let x_1, x_2 be any two points of $[-1, 1]$, then

$$|f(x_1) - f(x_2)| = |x_1^2 - x_2^2| = |x_1 - x_2| \cdot |x_1 + x_2|$$

$$< \varepsilon, \text{ when } |x_1 - x_2| < \frac{1}{2}\varepsilon = \delta$$

(where δ is independent of the choice of x_1, x_2).

Thus, for any $\varepsilon > 0, \exists$ a $\delta = \frac{1}{2}\varepsilon$ such that for any choice of x_1, x_2 in [–1, 1], we have

$$|f(x_1) - f(x_2)| < \varepsilon, \text{ when } |x_1 - x_2| < \frac{1}{2}\varepsilon = \delta$$

Thus, the function f is uniformly continuous on [–1, 1].

Example 28. Prove that $\sin x$ is uniformly continuous on $[0, \infty[$.

■ Let $\varepsilon > 0$ be given. Let $x, y \in [0, \infty[$, then

$$|\sin x - \sin y| = \left|2\sin\left(\frac{x-y}{2}\right)\cos\left(\frac{x+y}{2}\right)\right| \le 2\frac{|x-y|}{2}$$

$$= |x - y| < \varepsilon, \text{ whenever } |x - y| < \varepsilon$$

Therefore, taking $\delta = \varepsilon$

$|\sin x - \sin y| < \varepsilon$, whenever $|x - y| < \delta$

Hence, $\sin x$ is uniformly continuous on $[0, \infty[$.

EXERCISE

1. Show that the following functions are uniformly continuous in the given interval:
   (i) $f(x) = x^2$ in [1, 2]
   (ii) $f(x) = x^3$ in [0, 1]
   (iii) $f(x) = \sqrt{x}$ in $[0, 2]$
   (iv) $f(x) = x/(1 + x^2)$ on **R**.
2. Show that the function $f(x) = x^2$ is not uniformly continuous on $[0, \infty[$.
3. Show that the function $f(x) = 1/x^2$ is uniformly continuous on $[a, \infty[$, where $a > 0$; but not uniformly continuous on $]0, \infty[$.
4. Determine which of the following functions are uniformly continuous on the indicated intervals:
   (i) $f(x) = x^3$ on $[0, \infty[$
   (ii) $f(x) = \tan^{-1} x$ on **R**
   (iii) $f(x) = \dfrac{\sin x}{x}$ on $]0, \infty[$
   (iv) $f(x) = \begin{cases} \sin \pi x & \text{for } x \in]0, 1] \\ x^2 - 1 & \text{for } x \in]1, 2[. \end{cases}$

5. If f and g are uniformly continuous on the same interval, prove that $f+g$ and $f-g$ are also uniformly continuous on the interval.
6. Prove that if f and g are each uniformly continuous on the bounded open interval $]a, b[$, then the product fg is uniformly continuous on $]a, b[$.
7. Prove that if f and g are each uniformly continuous on the interval I and if in addition each function is bounded on I, then the product fg is uniformly continuous on I. Is boundedness of each function on I necessary for the uniform continuity of the product?

   [*Hint*: Boundedness of each function on I is not necessary; consider $f(x) = g(x) = \sqrt{x}$ on $[0, \infty[$].
8. Show by an example that a continuous bounded function on the bounded open interval $]a, b[$ need not be uniformly continuous on $]a, b[$.
9. Prove or give a counter example:

   If $f(x)$ is continuous and bounded on **R**, then f is uniformly continuous on **R**.

   [*Hint*: False, $f(x) = \sin x^2,\ x \in \mathbf{R}$]
10. If f is continuous on bounded open interval $]a, b[$, then prove that f is uniformly continuous on $]a, b[$ iff $f(a+)$ and $f(b-)$ both exist.
11. If f is continuous on $[a, \infty[(\text{or}]-\infty, b])$ and $\lim\limits_{x\to\infty} f(x) \left(\text{or } \lim\limits_{x\to-\infty} f(x)\right)$ exists, then prove that f is uniformly continuous on $[a, \infty[(\text{or}]-\infty, b])$. Is the converse true?
12. Prove that, if f is continuous on **R**, then f is uniformly continuous on every bounded interval of **R**. Is f then uniformly continuous on **R**?
13. If f is continuous on **R** and $\lim\limits_{x\to\infty} f(x)$ and $\lim\limits_{x\to-\infty} f(x)$ both exist, then prove that f is uniformly continuous on **R**.

ANSWERS

4. (*i*) No (*ii*) Yes (*iii*) Yes (*iv*) Yes.

6 Functions of a Single Variable (II)

1. THE DERIVATIVE

In this chapter we shall study the derivative, its existence and applications. We shall be concerned mainly with the real valued functions of a real variable, *i.e.*, the domains and the ranges of the functions considered here will be sets of real numbers.

1.1 Derivability at a Point

Let f be a real valued function defined on an interval $I = [a, b] \subseteq \mathbf{R}$. It is said to be derivable at an *interior point* c (where $a < c < b$) if

$$\lim_{h \to 0} \frac{f(c+h) - f(c)}{h} \text{ or } \lim_{x \to c} \frac{f(x) - f(c)}{x - c} \text{ exists.}$$

The limit in case it exists, is called the *Derivative* or the *Differential Coefficient* of the function f at $x = c$, and is denoted by $f'(c)$. The limit exists when the left-hand and the right-hand limits exist and are equal.

$\lim\limits_{x \to c-0} \dfrac{f(x) - f(c)}{x - c}$ is called the *Left-hand Derivative* and is denoted by

$$f'(c-0),\ f'(c-) \text{ or } Lf'(c),$$

while $\lim\limits_{x \to c+0} \dfrac{f(x) - f(c)}{x - c}$ is called the *Right-hand Derivative* and is denoted by

$$f'(c+0),\ f'(c+) \text{ or } Rf'(c).$$

Thus, the derivative $f'(c)$ exist when

$$Lf'(c) = Rf'(c)$$

1.2 Derivability in an Interval

A function f defined on $[a, b]$ is *derivable at the end point a, i.e.,* $f'(a)$ exists if,

$$\lim_{x \to a+0} \frac{f(x) - f(a)}{x - a} \text{ exists.}$$

In other words,

$$f'(a) = \lim_{x \to a+0} \frac{f(x) - f(a)}{x - a}$$

Similarly, it is *derivable at the end point b,* if $\lim\limits_{x \to b-0} \dfrac{f(x) - f(b)}{x - b}$ exists.

If a function is derivable at all points of an interval except the end points, it is said to be derivable in the *open interval.*

A function is derivable in the *closed interval* $[a, b]$, if it is derivable in the open interval $]a, b[$ and also at the end points a and b.

If f is not differentiable at $x = a$, then the upper and lower one-sided limits

$x \to a+$ or $x \to a-$ of $\dfrac{f(x) - f(a)}{x - a}$ will exist (possibly ∞ or $-\infty$).

These are denoted by $D^+ f, D_+ f, D^- f, D_- f$ respectively, and are called the *Dini* derivatives at a. For example,

$$D^+ f(a) = \overline{\lim_{x \to a+}} \frac{f(x) - f(a)}{x - a} = \overline{\lim_{h \to 0+}} \frac{f(a+h) - f(a)}{h}$$

$$D^+ f(a) = \underline{\lim}_{x \to a+} \frac{f(x) - f(a)}{x - a} = \underline{\lim}_{h \to 0+} \frac{f(a+h) - f(a)}{h}$$

If $D^+ f(a) = D_+ f(a) = l$, then we say that the right hand derivative at a exists and its value is l. Similarly, if $D^- f(a) = D_- f(a) = m$, then we say that the left hand derivative exists and is equal to m, where

$$\overline{\lim_{x \to a}}\, f(x) = \inf_{\delta > 0} \sup \{f(x) : 0 < |x - a| < \delta\}, \text{ and}$$

$$\underline{\lim}_{x \to a}\, f(x) = \sup_{\delta > 0} \inf \{f(x) : 0 < |x - a| < \delta\}$$

for f to be defined, bounded and real in

$$]a - \delta, a + \delta[, \delta > 0$$

If f is unbounded above in $]a - \delta, a + \delta[$, then $\overline{\lim}\, f(x) = +\infty$ and for f to be unbounded below

$$\underline{\lim}\, f(x) = -\infty.$$

Example 1. Show that the function $f(x) = x^2$ is derivable on $[0, 1]$.

■ Let x_0 be any point of $]0, 1[$, then

$$f'(x_0) = \lim_{x \to x_0} \frac{x^2 - x_0^2}{x - x_0} = \lim_{x \to x_0} (x + x_0) = 2x_0$$

At the end points, we have

$$f'(0) = \lim_{x\to 0+0} \frac{f(x) - f(0)}{x - 0} = \lim_{x\to 0+} \frac{x^2}{x} = \lim_{x\to 0+} x = 0$$

$$f'(1) = \lim_{x\to 1-0} \frac{f(x) - f(1)}{x - 1} = \lim_{x\to 1-} \frac{x^2 - 1}{x - 1} = \lim_{x\to 1-} (x + 1) = 2$$

Thus, the function is derivable in the closed interval [0, 1].

Example 2. A function f is defined on **R** by

$$f(x) = x \quad \text{if } 0 \le x < 1$$
$$= 1 \quad \text{if } x \ge 1$$

■ Consider the derivability at $x = 1$.

$$Lf'(1) = \lim_{x\to 1-} \frac{f(x) - f(1)}{x - 1} = \lim_{x\to 1-} \frac{x - 1}{x - 1} = 1$$

$$Rf'(1) = \lim_{x\to 1+} \frac{f(x) - f(1)}{x - 1} = \lim_{x\to 1+} \frac{1 - 1}{x - 1} = 0$$

$\therefore$ $$Lf'(1) \ne Rf'(1)$$

Thus, $\lim_{x\to 1} \dfrac{f(x) - f(1)}{x - 1}$ does not exist, *i.e.*, $f'(1)$ does not exist.

Example 3. Consider the derivability of the function $f(x) = |x|$ at the origin.

■ $$\text{Left hand derivative} = \lim_{x\to 0-} \frac{f(x) - f(0)}{x - 0}$$

$$= \lim_{x\to 0-} \frac{|x|}{x} = \lim_{x\to 0-} \frac{-x}{x} = -1$$

$$\text{Right hand derivative} = \lim_{x\to 0+} \frac{f(x) - f(0)}{x - 0}$$

$$= \lim_{x\to 0+} \frac{|x|}{x} = \lim_{x\to 0+} \frac{x}{x} = 1$$

Thus,

$$f'(0-) \ne f'(0+)$$

Hence, the function is not derivable at $x = 0$.

Example 4. A function f defined as :

$$f(x) = \begin{cases} x^2 \sin(1/x) & \text{if } x \ne 0 \\ 0 & \text{if } x = 0 \end{cases}$$

is derivable at $x = 0$ but $\lim_{x\to 0} f'(x) \ne f'(0)$.

$$f'(0) = \lim_{x\to 0} \frac{f(x) - f(0)}{x - 0} = \lim_{x\to 0} \frac{x^2 \sin 1/x}{x}$$

$$= \lim_{x\to 0} (x \sin 1/x) = 0$$

From elementary calculus we know that for $x \neq 0$,

$$f'(x) = 2x \sin (1/x) - \cos (1/x)$$

Clearly $\lim_{x\to 0} f'(x)$ does not exist and therefore, there is no possibility of $\lim_{x\to 0} f'(x)$ being equal to $f'(0)$.

Thus, $f'(x)$ is not continuous at $x = 0$ but $f'(0)$ exists.

2. CONTINUOUS FUNCTIONS

In this section we shall consider a relation between derivability and continuity, viz.,

derivability at a point $\Rightarrow$ *continuity at that point*

Thus, we shall prove that continuity at a point is a necessary condition for derivability at that point.

Theorem 1. *A function which is derivable at a point is necessarily continuous at that point.*

Let a function f be derivable at $x = c$.

Hence, $f'(c) = \lim_{x\to c} \frac{f(x) - f(c)}{x - c}$ exists.

Now

$$f(x) - f(c) = \frac{f(x) - f(c)}{(x - c)} (x-c), \ (x \neq c)$$

Taking limits as $x \to c$, we have

$$\lim_{x\to c} \{f(x) - f(c)\} = \lim_{x\to c} \left\{\frac{f(x) - f(c)}{x - c} (x - c)\right\}$$

$$= \lim_{x\to c} \left\{\frac{f(x) - f(c)}{x - c}\right\} \cdot \lim_{x\to c} (x - c)$$

$$= f'(c) \cdot 0 = 0$$

so that $\lim_{x\to c} f(x) = f(c)$, and therefore, f is continuous at $x = c$.

2.1 It is to be clearly understood that while continuity is a necessary condition for derivability at a point, it is not a sufficient condition. We come across functions which are continuous at a point without being derivable there at, and still many more functions may be constructed.

Consider the function f defined by

$$f(x) = |x|, \ \forall\, x \in \mathbf{R}$$

$f(x)$ is continuous at $x = 0$, for

$$\lim_{x \to 0+} f(x) = \lim_{x \to 0-} f(x) = 0 = f(0)$$

But as shown in example 3, $f'(0)$ does not exist. Thus, the function is continuous but not derivable at the origin.

However, it was the genius of German mathematician Weierstrass, who gave a function which is continuous everywhere but not derivable anywhere, viz.,

$$f(x) = \sum_{n=1}^{\infty} \frac{1}{2^n} \cos(3^n x), \quad \forall\, x \in \mathbf{R}$$

Proof, however, is beyond the scope of the present book.

Some Counter Examples

1. $f(x) = |x| + |x - 1|, \quad \forall\, x \in \mathbf{R}$

   Continuous but not derivable at $x = 0$ and $x = 1$.

2. $f(x) = |x - \alpha|$

   Continuous but not derivable at $x = \alpha$.

3. $f(x) = x \sin 1/x \quad$ if $x \neq 0$

   $\quad = 0 \quad$ if $x = 0$

   Continuous but not derivable at the origin.

4. $f(x) = 0$ if $x \leq 0$

   $\quad = x$ if $x > 0$

   Continuous but not derivable at $x = 0$.

2.2 The existence of the derivative of a function at a point depends on the existence of a limit, viz., $\lim_{x \to c} \dfrac{f(x) - f(c)}{x - c}$. Therefore, keeping in view the corresponding theorems on limits, one can easily establish the following *fundamental theorem on derivatives.*

If the functions f, g are derivable at c, then the functions $f + g$, $f - g$, $f \cdot g$ *and* f/g $(g(c) \neq 0)$ *are also derivable at c, and*

$$(f \pm g)'(c) = f'(c) \pm g'(c)$$

$$(f \cdot g)'(c) = f'(c)g(c) + f(c)g'(c)$$

$$(f/g)'(c) = \{f'(c)g(c) - f(c)g'(c)\}/g''(c), \text{ if } g(c) \neq 0$$

To illustrate the procedure, we prove the following theorem.

Theorem 2. *If f is derivable at c and* $f(c) \neq 0$ *then the function* $1/f$ *is also derivable there at, and*

$$(1/f)'(c) = -f'(c)/\{f(c)\}^2$$

Since f is derivable at c, it is also continuous there at. Again since $f(c) \neq 0$, there exists a neighbourhood N of c wherein f does not vanish.

Now

$$\frac{1/f(x) - 1/f(c)}{x - c} = -\frac{f(x) - f(c)}{x - c} \cdot \frac{-1}{f(x)\,f(c)}, \quad x \in N$$

Proceeding to limits when $x \to c$, we get

$$\left(\frac{1}{f}\right)'(c) = \lim_{x \to c} \frac{1/f(x) - 1/f(c)}{x - c}$$

$$= -f'(c) \cdot \frac{1}{f(c) \cdot f(c)} = -\frac{f'(c)}{\{f(c)\}^2}$$

Thus, the limits exists and are equals, $-f'(c)/\{f(c)\}^2$.

Note: If f and g be two functions having the same domain D and $f \pm g$ or $f.g$ be derivable at $c \in D$, then f and g are not necessarily derivable at c. Consider, for instance

(*i*) $f(x) = \begin{cases} 1/x & x \neq 0 \\ \text{and } g(x) = -f(x) \\ 0 & x = 0 \end{cases}$

$f + g$ is derivable at the origin but f and g are not.

(*ii*) $f(x) = \begin{cases} x \sin(1/x) & x \neq 0 \\ \text{and } g(x) = x \\ 0 & x = 0 \end{cases}$

Here fg is derivable at the origin, whereas f is not.

(*iii*) $f(x) = |x|$ and $g(x) = -|x|$

$f(x) \cdot g(x) = -x^2$, so that $f \cdot g$ is derivable at the origin but the functions f and g are not derivable. Similarly many more functions may be constructed.

EXERCISE

1. Find the derivatives of the following functions at the indicated points:

   (*i*) $f(x) = K$, a constant, at $x = c$

   (*ii*) $f(x) = x$ at $x = 0$

   (*iii*) $f(x) = \sqrt{x}$ at $x = 4$

   (*iv*) $f(x) = e^x$ at $x = x_0$

2. Show that the function

   $$f(x) = |x| + |x - 1|$$

   is derivable at all points except 0 and 1.

3. $$f(x) = \begin{cases} x^3 \sin 1/x, & x \neq 0 \\ 0, & x = 0 \end{cases}$$

Prove that $f(x)$ has a derivative at $x = 0$ and that $f(x)$ and $f'(x)$ are continuous at $x = 0$.

4. $$f(x) = (x-a)\sin\frac{1}{x-a}, \quad x \neq a$$
$$= 0, \quad x = a$$

Show that $f(x)$ is continuous but not derivable at $x = a$.

5. Discuss the derivability of the following functions:

(i) $f(x) = \begin{cases} 2, & x \le 1 \\ x, & x > 1 \end{cases}$ at $x = 1$

(ii) $f(x) = \begin{cases} 1, & 0 \le x \le 1 \\ x, & x > 1 \end{cases}$ at $x = 1$

(iii) $f(x) = \begin{cases} 2x - 3, & 0 \le x \le 2 \\ x^2 - 3, & 2 < x \le 4 \end{cases}$ at $x = 2, 4$

6. Show that the function $f(x) = x\,|x|$ is derivable at the origin.

7. Find the derivative of f at $x = 0$, where $f(x) = x^2\,|x|$.

8. $f(x) = |x|$ and $g(x) = 3|x|$ $x \in R$, show that f, g is not derivable at the origin but $\lim\left[f(x)/g(x)\right]$ exists and is equal to $\lim\left[f'(x)/g'(x)\right]$, when $x \to 0$.

9. Find $Lf'(0)$ and $Rf'(0)$ for the following functions:

(i) $f(x) = \begin{cases} x \tan^{-1} 1/x, & x \neq 0 \\ 0, & x = 0 \end{cases}$

(ii) $f(x) = \begin{cases} \dfrac{x(e^{1/x} - 1)}{(e^{1/x} + 1)}, & x \neq 0 \\ 0, & x = 0 \end{cases}$

10. If $\quad f(x) = \begin{cases} \dfrac{x(e^{1/x} - e^{-1/x})}{(e^{1/x} + e^{-1/x})}, & x \neq 0 \\ 0, & x = 0 \end{cases}$

Show that f is continuous but not derivable at $x = 0$ and $Lf'(0) = -1$, $Rf'(0) = 1$.

11. Examine the function f, where $f(x) = \begin{cases} x^m \sin 1/x, & x \neq \theta \\ 0, & x = 0 \end{cases}$ for derivability at the origin. Also determine m where f' is continuous at the origin.

12. If functions f and g are defined on $[0, \infty[$ by

$$f(x) = \lim_{n\to\infty} \frac{x^n - 1}{x^n + 1} \text{ and } g(x) = \int_0^x f(t)\, dt\,;$$

then prove that g is continuous but not derivable at $x = 1$.

ANSWERS

5. (*i*) Not derivable (*ii*) Not derivable, $Lf'(1) = 0$, $Rf'(1) = 1$.

   (*iii*) Not derivable ; at $x = 2$, $f'(4) = 8$.

7. 0.

9. (*i*) $-\pi/2, \pi/2$ (*ii*) $-1, 1$.

11. Derivable if $m > 1$; $m > 2$.

3. INCREASING AND DECREASING FUNCTIONS

A function f is said to be increasing or decreasing at a point $x = c$ according as the value of the function increases or decreases at that point with increase in x. Thus for any x in the neighbourhood $]c - \delta, c+\delta[, \delta>0$ of c,

$$f(c - \delta) \le f(x) \le f(c + \delta) \text{ for an increasing function.}$$
$$f(c - \delta) \ge f(x) \ge f(c + \delta) \text{ for a decreasing function.}$$

A function is increasing or decreasing in the interval $[a, b]$ according as

$$f(x_2) \ge f(x_1) \text{ or } f(x_2) \le f(x_1),\ \forall\, x_2 \ge x_1 \wedge x_1, x_2 \in [a, b].$$

The function is *strictly increasing* or *strictly decreasing* if the strict inequality holds in the above relations, *i.e.*,

$$f(x_2) > f(x_1) \text{ for a strictly increasing function,}$$
$$f(x_2) < f(x_1) \text{ for a strictly decreasing function}$$
$$\forall\, x_2 > x_1 \wedge x_1, x_2 \in [a, b]$$

A function is said to be *monotone* or *monotonic* in an interval I if it is either increasing in I or decreasing in I.

It is said to be *strictly monotone* in I if it is either strictly increasing or strictly decreasing in I.

3.1 Sign of the Derivative

Let c *be an interior point* of the domain $[a, b]$ of a function f and let $f'(c)$ exist and be positive.

By definition,

$$\lim_{x\to c} \frac{f(x) - f(c)}{x - c} = f'(c),\ x \ne c$$

Thus depending on the choice of a positive ε however small, $\exists$ a positive number δ such that

$$\left| \frac{f(x) - f(c)}{x - c} - f'(c) \right| < \varepsilon, \text{ when } |x - c| < \delta,\ x \ne c$$

or

$$f'(c) - \varepsilon < \frac{f(x) - f(c)}{x - c} < f'(c) + \varepsilon, \text{ when } x \in]c-\delta, c+\delta[, x \neq c$$

If $\varepsilon > 0$ is selected less than $f'(c)$, $\dfrac{f(x) - f(c)}{x - c}$ lies between two positive numbers and is therefore itself positive.

Hence,

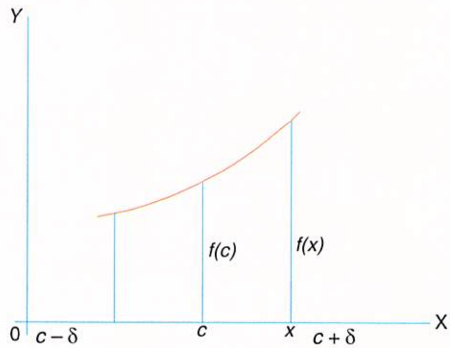
$$\frac{f(x) - f(c)}{x - c} > 0, \text{ when } x \in]c - \delta, c+\delta[, x \neq c.$$

Thus

(*i*) $f(x) - f(c) > 0$ or $f(x) > f(c)$ and $x \in]c, c+\delta[$, and

(*ii*) $f(x) < f(c)$, when $x \in]c - \delta, c[$.

From (*i*) and (*ii*), we see that $f(x)$ is increasing at $x = c$.



Hence, the function is increasing at c if $f'(c) > 0$.

Similarly, it can be shown that the function is decreasing at $x = c$ if $f'(c) < 0$.

Let us now consider the *end points.*

(*a*) At the *end point* a, $\exists$ an interval $[a, a + \delta[$, such that

$$f'(a) > 0 \Rightarrow f(x) > f(a), \text{ for } x \in]a, a + \delta[$$
$$f'(a) < 0 \Rightarrow f(x) < f(a), \text{ for } x \in]a, a + \delta[$$

(*b*) At the *end point* b, $\exists$ an interval $]b - \delta, b]$, such that

$$f'(b) > 0 \Rightarrow f(x) < f(b), \text{ for } x \in]b - \delta, b[$$
$$f'(b) < 0 \Rightarrow f(x) > f(b), \text{ for } x \in]b - \delta, b[$$

Example 5. Show that log $(1 + x)$ lies between

$$x - \frac{x^2}{2} \text{ and } x - \frac{x^2}{2(1+x)}, \quad \forall\, x > 0$$

■ Consider

$$f(x) = \log(1+x) - \left(x - \frac{x^2}{2}\right)$$

$$\therefore \qquad f'(x) = \frac{1}{1+x} - (1-x) = \frac{x^2}{1+x} > 0, \ \forall\, x > 0$$

Hence, $f(x)$ is an increasing function for all $x > 0$.

Also

$$f(0) = 0$$

Hence for $x > 0$, $f(x) > 0$.

Thus

$$\log(1+x) > x - \frac{x^2}{2}, \quad \text{for } x > 0$$

Similarly by considering the function

$$F(x) = x - \frac{x^2}{2(1+x)} - \log(1+x),$$

it can be shown that

$$\log(1+x) < x - \frac{x^2}{2(1+x)}, \quad \forall\, x > 0.$$

Ex. 1. Show that

$$x^3 - 6x^2 + 15x + 3 > 0, \quad \forall\, x > 0$$

Ex. 2. Show that

(*i*) $\dfrac{x}{1+x} < \log(1+x) < x, \ \forall\, x > 0$

(*ii*) $\dfrac{x}{1+x^2} < \tan^{-1} x < x, \quad \forall\, x > 0$

Ex. 3. Show that

(*i*) $\tan x > x, 0 < x < \pi/2$

(*ii*) $\dfrac{2}{\pi} \le \dfrac{\sin x}{x} < 1, 0 < |x| \le \pi/2.$

Ex. 4. Show that

$$2x < \log\frac{1+x}{1-x} < 2x\left(1 + \frac{x^2}{3(1-x^2)}\right), 0 < x < 1.$$

Ex. 5. Show that

$$2/(2x+1) < \log(1+1/x) < 1/\sqrt{x(x+1)}, \quad \forall\, x > 0.$$

4. DARBOUX'S THEOREM

If a function f is derivable on a closed interval $[a, b]$ and $f'(a)$, $f'(b)$ are of opposite signs then there exists at least one point c between a and b such that $f'(c) = 0$.

For the sake of definiteness, let us take $f'(a) < 0$ and $f'(b) > 0$.

Since $f'(a)$ is negative, therefore, there exists a positive number δ_1 such that

$$f(x) < f(a), \quad \forall\, x \in]a, a+\delta_1[\qquad ...(1)$$

Again, since $f'(b)$ is positive, there exists a positive number δ_2 such that

$$f(x) < f(b), \quad \forall\, x \in]b-\delta_2, b[\qquad ...(2)$$

Also, since f is derivable in $[a, b]$, it is continuous in the closed interval $[a, b]$. Being continuous in the closed interval, it is bounded and attains its bounds. Thus if m is the infimum (*g. l. b.*), $\exists$ a point $c \in [a, b]$ such that

$$f(c) = m$$

It is clear from (1) and (2) that c cannot coincide with a or b for otherwise there would exist points where the values of the function would be less than $f(c)$. Thus, c is an interior point of $[a, b]$.

Now, we proceed to show that it is this point where $f'(c) = 0$.

If $f'(c) < 0$, then $\exists$ an interval $]c, c+\delta_3[$, $\delta_3 > 0$ such that $\forall\, x \in]c, c+\delta_3[$, $f(x) < f(c) = m$, which contradicts the fact that m is infimum of f.

Again, if $f'(c) > 0$, $\exists$ an interval $]c-\delta_4, c[$, $\delta_4 > 0$ for every point x of which

$f(x) < f(c) = m$, which again is a contradiction.

Hence, the only possibility, $f'(c) = 0$.

Note: If $f'(a) > 0$ and $f'(b) < 0$, then proceeding as above, it can be shown that $f'(d) = 0$ where $d \in]a, b[$ is the point where the function attains the supremum.

Intermediate value theorem for derivatives. *If a function f is derivable on a closed interval $[a, b]$ and $f'(a) \neq f'(b)$ and k, a number lying between $f'(a)$ and $f'(b)$ then $\exists$ at least one point $c \in]a, b[$ such that $f'(c) = k$.*

Let $g(x) = f(x) - kx$, $x \in [a, b]$.

Clearly $g(x)$ is derivable on $[a, b]$ and

$$g'(a) = f'(a) - k, \; g'(b) = f'(b) - k$$

Since k lies between $f'(a)$ and $f'(b)$.

$\therefore$ $g'(a)$ and $g'(b)$ are of opposite signs.

$\therefore$ $g(x)$ satisfies the conditions of Darboux's theorem.

Thus there exists at least one point $c \in]a, b[$ such that

$$g'(c) = 0 \text{ or } f'(c) = k.$$

Note: The derivative f' (not necessarily continuous) satisfies the intermediate-value property, and so f' has no discontinuities of the first kind or removable. As a consequence, it follows that, *monotone derivatives are necessarily continuous.*

5. ROLLE'S THEOREM

If a function f defined on $[a, b]$ *is*

(*i*) *continuous on* $[a, b]$,

(*ii*) *derivable on* $]a, b[$, *and*

(*iii*) $f(a) = f(b)$,

then there exists at least one real number c between a and b $(a < c < b)$ *such that* $f'(c) = 0$.

Since the function is continuous on the closed interval $[a, b]$ it is bounded and attains its bounds. Thus, if m and M are the infimum (*g. l. b.*) and the supremum (*l.u.b.*) respectively of the function f, then $\exists$ points c and d of $[a, b]$ such that

$$f(c) = m \text{ and } f(d) = M.$$

There are two possibilities: either $m = M$ or $m \neq M$.

If $m = M$, then f is constant over $[a, b]$ and therefore, its derivatives

$$f'(x) = 0, \quad \forall\, x \in [a, b].$$

When $m \neq M$, both of these cannot be equal to the same quantity $f(a)$. At least one of these, say, m is different from $f(a)$ or $f(b)$, so that

$$f(c) = m \neq f(a) \Rightarrow c \neq a$$

$$f(c) = m \neq f(b) \Rightarrow c \neq b$$

This means that c lies in the open interval $]a, b[$.

We shall now show that c is the point where $f'(c) = 0$.

If $f'(c) < 0$, then $\exists$ an interval $]c, c + \delta_1[, \delta_1 > 0$ for every point x of which $f(x) < f(c) = m$; which contradicts the fact that m is the infimum.

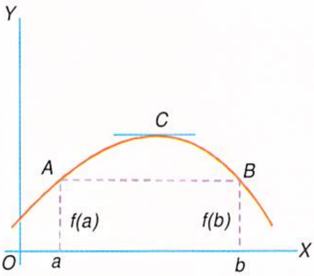
If $f'(c) > 0, \exists$ an interval $]c - \delta_2, c[, \delta_2 > 0$ for every point x of which $f(x) < f(c) = m$; which is also a contradiction.

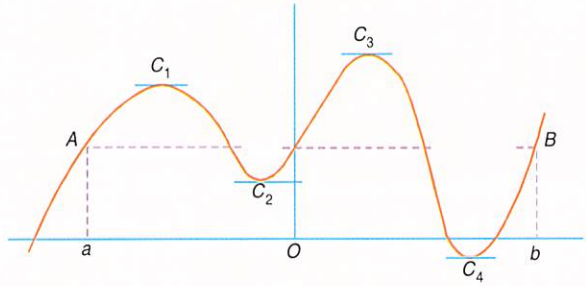
Hence, the only possibility is $f'(c) = 0$.

5.1 Interpretation of Rolle's Theorem

Geometric. Let the curve $y = f(x)$, which is continuous on $[a, b]$ and derivable on $]a, b[$ be drawn.

The theorem simply states that between two points with equal ordinates on the graph of f, there exists at least one point where the tangent is parallel to x-axis.





Algebraic. Between two zeros a and b of $f(x)$ (*i.e.*, between two roots a and b of $f(x) = 0$) there exists at least one zero of $f'(x)$.

Ex. 1. Show that between two consecutive zeros of $f'(x)$ there lies at the most one zero of $f(x)$.

Ex. 2. Show that, for any real number k, the polynomial

$f(x) = x^3 + x + k$ has exactly one real root.

6. LAGRANGE'S MEAN VALUE THEOREM
(First mean value theorem of differential calculus)

If a function f defined on $[a, b]$ is

(*i*) *continuous on $[a, b]$, and*

(*ii*) *derivable on $]a, b[$,*

then there exists at least one real number c between a and b ($c \in]a, b[$) such that

$$\frac{f(b) - f(a)}{b - a} = f'(c)$$

Let us consider a function

$$\phi(x) = f(x) + Ax,\ x \in [a, b],$$

where A is a constant to be determined such that $\phi(a) = \phi(b)$.

$$\therefore \qquad A = -\frac{f(b) - f(a)}{b - a}$$

Now the function $\phi(x)$, being the sum of two continuous and derivable functions, is itself

(*i*) continuous on $[a, b]$,

(*ii*) derivable on $]a, b[$, and

(*iii*) $\phi(a) = \phi(b)$.

Therefore, by Rolle's Theorem $\exists$ a real number $c \in]a, b[$ such that $\phi'(c) = 0$.

But

$$\phi'(x) = f'(x) + A$$

$$\therefore \qquad 0 = \phi'(c) = f'(c) + A$$

or

$$f'(c) = -A = \frac{f(b) - f(a)}{b - a}$$

Another Statement. If in the statement of the theorem, b is replaced by $a + h$, then the number c between a and b may be written as $a + \theta h$, where $0 < \theta < 1$. Thus,

$$f(a + h) - f(a) = hf'(a + \theta h)$$

or

$$f(a + h) = f(a) + hf'(a + \theta h), \text{ where } 0 < \theta < 1.$$

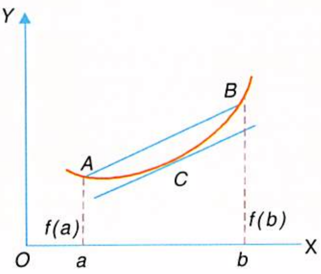
6.1 Deductions

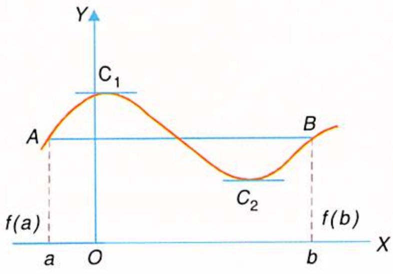
1. If a function $f(x)$ satisfies the conditions of the Mean Value Theorem and $f'(x) = 0$ for all $x \in]a, b[$, then $f(x)$ is constant on $[a, b]$.
   Let x_1, x_2 (where $x_1 < x_2$) be any two distinct points of $[a, b]$.
   Hence by Lagrange's Mean Value Theorem,
   $$\frac{f(x_2) - f(x_1)}{x_2 - x_1} = f'(c) = 0, \text{ where } x_1 < c < x_2$$
   Thus,
   $$f(x_2) = f(x_1)$$
   Hence, the function keeps the same value and is therefore constant on $[a, b]$.
2. If two functions have equal derivatives at all points of $]a, b[$, then they differ only by a constant.
3. If a function f is (*i*) continuous on $[a, b]$, (*ii*) derivable on $]a, b[$, and (*iii*) $f'(x) > 0, \ \forall\, x \in]a, b[$, then f is strictly increasing on $[a, b]$.

   Let x_1, x_2 (where $x_1 < x_2$) be any two distinct points of $[a, b]$, then by Lagrange's Mean Value Theorem
   $$\frac{f(x_2) - f(x_1)}{x_2 - x_1} = f'(c) > 0, \text{ for } x_1 < c < x_2$$
   or
   $$f(x_2) - f(x_1) > 0 \Rightarrow f(x_2) > f(x_1), \text{ for } x_2 > x_1$$
   Thus, f is strictly increasing on $[a, b]$.
4. If f' exists and is bounded on some interval I, then f is uniformly continuous on I.

6.2 Geometrical Interpretation

The theorem simply states that between two points A and B of the graph of f there exists at least one point where the tangent is parallel to the chord AB.





7. CAUCHY'S MEAN VALUE THEOREM
(Second mean value theorem)

If two functions f, g defined on $[a, b]$ *are*

(*i*) *continuous on* $[a, b]$,

(*ii*) *derivable on* $]a, b[$, *and*

(*iii*) $g'(x) \neq 0$, *for any* $x \in]a, b[$,

then there exists at least one real number c between a and b $(c \in]a, b[)$ *such that*

$$\frac{f(b)-f(a)}{g(b)-g(a)} = \frac{f'(c)}{g'(c)}$$

At the outset we notice that $g(a) \neq g(b)$, for otherwise, in view of the conditions (*i*) and (*ii*) of the theorem, $g(x)$ would satisfy all the conditions of Rolle's Theorem and its derivative $g'(x)$ would vanish for some $x \in]a, b[$ contrary to (*iii*).

Consider the function

$$\phi(x) = f(x) + Ag(x),\ x \in [a, b],$$

where A is a constant to be determined such that $\phi(a) = \phi(b)$.

$$\therefore \qquad A = -\frac{f(b)-f(a)}{g(b)-g(a)}$$

Now the function $\phi(x)$, being the sum of two continuous and derivable functions, is itself:

(*i*) *continuous on* $[a, b]$,

(*ii*) *derivable on* $]a, b[$, and

(*iii*) $\phi(a) = \phi(b)$.

Therefore, by Rolle's Theorem $\exists$ a real number $c \in]a, b[$ such that $\phi'(c) = 0$

But $\phi'(x) = f'(x) + Ag'(x)$

$$\Rightarrow \qquad 0 = \phi'(c) = f'(c) + Ag'(c)$$

$$\therefore \qquad \frac{f'(c)}{g'(c)} = -A = \frac{f(b)-f(a)}{g(b)-g(a)}.$$

Another Statement. If two functions f, g defined on $[a, a+h]$ are continuous on $[a, a+h]$, derivable on $]a, a+h[$ and $g'(x) \neq 0$ for any $x \in]a, a+h[$, then there exists at least one real number θ between 0 and 1 such that

$$\frac{f(a+h) - f(a)}{g(a+h) - g(a)} = \frac{f'(a+\theta h)}{g'(a+\theta h)}, \quad 0 < \theta < 1.$$

Corollary. *Lagrange's Mean Value Theorem* may be deduced as a particular case, for $g(x) = x$.

Note: Cauchy's Mean Value Theorem cannot be deduced by applying Lagrange's Mean Value Theorem separately to the two functions and then dividing, for, then we get

$$\frac{f(b) - f(a)}{g(b) - g(a)} = \frac{f'(c_1)}{g'(c_2)}$$

where c_1 and c_2 may not be equal.

Geometrical (Physical) Interpretation. We may write

$$\frac{\{f(b) - f(a)\}/(b-a)}{\{g(b) - g(a)\}/(b-a)} = \frac{f'(c)}{g'(c)}$$

Hence, the ratio of the mean rates of increase of two functions in an interval is equal to the ratio of the actual rates of increase of the functions at some point within the interval.

Example 6. Show that

$$\frac{v-u}{1+v^2} < \tan^{-1} v - \tan^{-1} u < \frac{v-u}{1+u^2}, \text{ if } 0 < u < v$$

and deduce that

$$\frac{\pi}{4} + \frac{3}{25} < \tan^{-1}\frac{4}{3} < \frac{\pi}{4} + \frac{1}{6}$$

■ Let $f(x) = \tan^{-1} x$, then for $u < x < v$,

$$f'(x) = \frac{1}{1+x^2}$$

Applying the Mean Value Theorem to f, we get

$$\frac{\tan^{-1} v - \tan^{-1} u}{v-u} = \frac{1}{1+\xi^2}, \text{ for } u < \xi < v$$

But

$$\xi > u \Rightarrow \frac{1}{1+\xi^2} < \frac{1}{1+u^2}$$

and

$$\xi < v \Rightarrow \frac{1}{1+\xi^2} > \frac{1}{1+v^2}$$

$\therefore$

$$\frac{1}{1+v^2} < \frac{\tan^{-1} v - \tan^{-1} u}{v-u} < \frac{1}{1+u^2}$$

or

$$\frac{v-u}{1+v^2} < \tan^{-1} v - \tan^{-1} u < \frac{v-u}{1+u^2}$$

The other result follows by taking $v = \frac{4}{3}$ and $u = 1$.

Example 7. Show that

$$\frac{\sin\alpha - \sin\beta}{\cos\beta - \cos\alpha} = \cot\theta, \text{ where } 0 < \alpha < \theta < \beta < \frac{\pi}{2}$$

■ Let $f(x) = \sin x$ and $g(x) = \cos x$, for $x \in [\alpha, \beta]$.

$\therefore$ $$f'(x) = \cos x \text{ and } g'(x) = -\sin x$$

Functions f and g are both continuous and differentiable, therefore by Cauchy's Mean Value Theorem on $[\alpha, \beta]$,

$$\frac{\sin\beta - \sin\alpha}{\cos\beta - \cos\alpha} = \frac{\cos\theta}{-\sin\theta}, \alpha < \theta < \beta$$

or

$$\frac{\sin\alpha - \sin\beta}{\cos\beta - \cos\alpha} = \cot\theta, \alpha < \theta < \beta.$$

Example 8. A twice differentiable function f is such that $f(a) = f(b) = 0$ and $f(c) > 0$, for $a < c < b$. Prove that there is at least one value ξ between a and b for which $f''(\xi) < 0$.

■ Let us consider the function f on $[a, b]$.

Since f'' exists, f' and f both exist and are continuous on $[a, b]$. Since c is a point between a and b, applying Lagrange's Mean Value Theorem to f on the intervals $[a, c]$ and $[c, b]$ respectively, we get

$$\frac{f(c) - f(a)}{c - a} = f'(\xi_1), \quad a < \xi_1 < c$$

and

$$\frac{f(b) - f(c)}{b - c} = f'(\xi_2), \quad c < \xi_2 < b$$

But

$$f(a) = f(b) = 0$$

$\therefore$ $$f'(\xi_1) = \frac{f(c)}{c-a} \text{ and } f'(\xi_2) = -\frac{f(c)}{b-c}$$

where $a < \xi_1 < c < \xi_2 < b$.

Again $f'(x)$ is continuous and derivable on $[\xi_1, \xi_2]$. Therefore, by Mean Value Theorem,

$$\frac{f'(\xi_2) - f'(\xi_1)}{\xi_2 - \xi_1} = f''(\xi), \text{ where } \xi_1 < \xi < \xi_2$$

Substituting the values of $f'(\xi_2)$ and $f'(\xi_1)$, we get

$$f''(\xi) = \frac{-f(c)}{\xi_2 - \xi_1}\left(\frac{1}{b-c} + \frac{1}{c-a}\right)$$

$$= \frac{-(b-a)f(c)}{(\xi_2 - \xi_1)(b-c)(c-a)} < 0.$$

Example 9. If a function f is such that its derivative f' is continuous on $[a, b]$ and derivable on $]a, b[$, then show that there exists a number c between a and b such that

$$f(b) = f(a) + (b-a)f'(a) + \frac{1}{2}(b-a)^2 f''(c)$$

■ Clearly the functions f and f' are continuous and derivable on $[a, b]$.

Consider the function

$$\phi(x) = f(b) - f(x) - (b-x)f'(x) - (b-x)^2 A$$

where A is a constant to be determined such that $\phi(a) = \phi(b)$.

$$\therefore \quad f(b) - f(a) - (b-a)f'(a) - (b-a)^2 A = 0 \qquad ...(1)$$

Now $\phi(x)$, being the sum of continuous and derivable functions, is itself continuous on $[a, b]$ and derivable on $]a, b[$ and also $\phi(a) = \phi(b)$.

Thus $\phi(x)$ satisfies all the conditions of Rolle's Theorem and therefore $\exists\, c \in]a, b[$ such that $\phi'(c) = 0$.

Now,

$$\phi'(x) = -f'(x) + f'(x) - (b-x)f''(x) + 2(b-x)A$$

$$\therefore \quad -(b-c)f''(c) + 2(b-c)A = \phi'(c) = 0$$

But

$$b - c \neq 0$$

$$\therefore \quad A = \frac{1}{2}f''(c) \qquad ...(2)$$

Hence, from (1) and (2)

$$f(b) = f(a) + (b-a)f'(a) + \frac{1}{2}(b-a)^2 f''(c)$$

Note: Motivation for $\phi(x)$ is attained by replacing a by x and transposing all the terms to the right in the result to be proved.

Example 10. Assuming f'' to be continuous on $[a, b]$, show that

$$f(c) - f(a)\frac{b-c}{b-a} - \frac{c-a}{b-a}f(b) = \frac{1}{2}(c-a)(c-b)f''(\xi)$$

where c and ξ both lie in $[a, b]$.

■ We have to show that

$$(b-a)f(c)-(b-c)f(a)-(c-a)\,f(b)=\frac{1}{2}(b-a)(c-a)(c-b)f''(\xi)$$

Consider the function, for $x \in [a, b]$ defined by

$$\phi(x) = (b-a)f(x) - (b-x)\,f(a) - (x-a)f(b) - (b-a)(x-a)(x-b)\,A$$

where A is a constant to be determined such that $\phi(c) = 0$.

$$\therefore \quad (b-a)f(c) - (b-c)\,f(a) - (c-a)f(b) - (b-a)(c-a)(c-b)\,A = 0 \quad \text{...(1)}$$

Clearly $\phi(a) = 0 = \phi(b)$, and $\phi(x)$ is differentiable in $[a, b]$.

The function ϕ satisfies all the conditions of Rolle's Theorem on each of the intervals $[a, c]$ and $[c, b]$ and therefore $\exists$ two numbers ξ_1, ξ_2 in $]a, c[$ and $]c, b[$ respectively, such that $\phi'(\xi_1) = 0$ and $\phi'(\xi_2) = 0$.

Again

$$\phi'(x) = (b-a)f'(x) + f(a) - f(b) - (b-a)\,\{2x - (a+b)\}\,A$$

which is continuous on $[a, b]$ and derivable on $]a, b[$ and in particular on $[\xi_1, \xi_2]$.

Also $\phi'(\xi_1) = \phi'(\xi_2) = 0$

Therefore by Rolle's Theorem $\exists\ \xi \in]\xi_1, \xi_2[$ such that $\phi''(\xi) = 0$.

Now $\phi''(x) = (b-a)\,f''(x) - 2(b-a)\,A$

so that

$$f''(\xi) - 2A = 0, b \neq a$$

or

$$A = \frac{1}{2}f''(\xi), \text{ where } a < \xi_1 < \xi < \xi_2 < b \quad \text{...(2)}$$

From (1) and (2), the result follows.

Example 11. Show that $\dfrac{\tan x}{x} > \dfrac{x}{\sin x}$, for $0 < x < \dfrac{\pi}{2}$

■ We have thus to show that

$$\frac{\tan x}{x} - \frac{x}{\sin x} > 0 \text{ or } \frac{\sin x \tan x - x^2}{x \sin x} > 0, \text{ for } 0 < x < \frac{\pi}{2}.$$

Since $x \sin x > 0$, for $0 < x < \pi/2$, it will therefore suffice to show that $\sin x \tan x - x^2 > 0$

Let $f(x) = \sin x \tan x - x^2$ then for $0 < x < \dfrac{\pi}{2}$,

$$f'(x) = \cos x \tan x + \sin x \sec^2 x - 2x = \sin x + \sin x \sec^2 x - 2x$$

We cannot decide about the sign of $f'(x)$ mainly because of the presence of the $2x$ term.

The function $f'(x)$ is continuous and derivable on $]0, \pi/2[$.

$$\therefore \quad f''(x) = \cos x + \cos x \sec^2 x + 2\sin x \sec^2 x \tan x - 2$$

$$= \left(\sqrt{\sec x} - \sqrt{\cos x}\right)^2 + 2\tan^2 x \sec x > 0, \text{ for } 0 < x < \pi/2$$

Since the derivative $f''(x)$ of $f'(x)$ is positive, the function $f'(x)$ is an increasing function. Further since $f'(0)=0$, therefore the function $f'(x)>0$ for $0<x<\pi/2$.

Again, since $f'(x)>0$, $f(x)$ is an increasing function and because $f(0)=0$, the function $f(x)>0$, for $0<x<\pi/2$.

Thus it follows that

$$\frac{\tan x}{x} > \frac{x}{\sin x}, \text{ for } 0<x<\pi/2.$$

Note: The above inequality can be put in the form:

$$\cos x < \left(\frac{\sin x}{x}\right)^2, 0<x<\frac{\pi}{2}.$$

Ex. Show that $\cos x < \left(\dfrac{\sin x}{x}\right)^3$, for $0<x<\dfrac{\pi}{2}$.

$$\left[\textit{Hint}: \text{Take } f(x) = x - \sin x \cos^{-1/3} x, 0 \le x < \frac{\pi}{2}\right].$$

EXERCISE

1. Examine the validity of the hypothesis and the conclusion of Rolle's Theorem:

   (*i*) $f(x) = x^3 - 4x$ on $[-2, 2]$,

   (*ii*) $f(x) = (x-a)^m (x-b)^n$, where m and n are positive integers on $[a, b]$,

   (*iii*) $f(x) = 1-(x-1)^{2/3}$ on $[0, 2]$,

   (*iv*) $f(x) = |x|$ on $[-1, 1]$,

   (*v*) $f(x) = 1-|x-1|$ on $[0, 2]$.

2. Examine the validity of the hypothesis and the conclusion of Lagrange's Mean Value Theorem:

   (*i*) $f(x) = |x|$ on $[-1, 1]$,

   (*ii*) $f(x) = \log x$ on $\left[\frac{1}{2}, 2\right]$,

   (*iii*) $f(x) = x(x-1)(x-2)$ on $\left[0, \frac{1}{2}\right]$,

   (*iv*) $f(x) = x^{1/3}$ on $[-1, 1]$,

   (*v*) $f(x) = 2x^2 - 7x + 10$ on $[2, 5]$.

3. Deduce Lagrange's Mean Value Theorem by considering the derivable function

$$\phi(x) = f(x) - f(a) - A(x-a), x \in [a, b].$$

4. Prove that between any two real roots of $e^x \sin x = 1$, there is at least one real root of $e^x \cos x + 1 = 0$.
[*Hint*: Apply Rolle's Theorem to the function $e^{-x} - \sin x$.]

5. A function f is continuous on $[a - h, a + h]$ and derivable on $]a - h, a + h[$, show that

$$f(a-h) - 2f(a) + f(a+h) = h[f'(a+\theta h) - f'(a-\theta h)], 0 < \theta < 1$$

[*Hint*: Use Mean Value Theorem for the function

$$\phi(t) = f(a+th) + f(a-ht) \text{ on } [0, 1].$$

6. If a function f is twice derivable on $[a, a + h]$, then show that

$$f(a+h) = f(a) + hf'(a) + \frac{h^2}{2} f''(a+\theta h), 0 < \theta < 1$$

7. If f', g' are continuous and differentiable on $[a, b]$, then show that for $a < c < b$.

$$\frac{f(b) - f(a) - (b-a)f'(a)}{g(b) - g(a) - (b-a)g'(a)} = \frac{f''(c)}{g''(c)}$$

[*Hint*: Apply Rolle's Theorem to the function

$$\phi(x) = f(x) + (b-x)f'(x) + A\{g(x) + (b-x)g'(x)\}]$$

8. If f', ϕ and ψ are continuous on $[a, b]$ and derivable on $]a, b[$, then show that there is a value c lying between a and b such that

$$\begin{vmatrix} f(a) & f(b) & f'(c) \\ \phi(a) & \phi(b) & \phi'(c) \\ \psi(a) & \psi(b) & \psi'(c) \end{vmatrix} = 0.$$

[*Hint*: Apply Rolle's Theorem to the function

$$g(x) = \begin{vmatrix} f(a) & f(b) & f(x) \\ \phi(a) & \phi(b) & \phi(x) \\ \psi(a) & \psi(b) & \psi(x) \end{vmatrix}].$$

9. A function f is such that its second derivative is continuous on $[a, a+h]$ and derivable on $]a, a+h[$. Show that there exists a number θ between 0 and 1 such that

$$f(a+h) - f(a) - \frac{1}{2}h\{f'(a) + f'(a+h)\} + \frac{h^3}{12}f''(a+\theta h) = 0.$$

10. If $f(0) = 0$ and $f''(x)$ exists for all $x > 0$, then show that

$$f'(x) - \frac{f(x)}{x} = \frac{1}{2}xf''(\xi), 0 < \xi < x.$$

Also deduce that if $f''(x)$ is positive for positive values of x, then $f(x)/x$ strictly increases as x increases.

11. If $f'(x)$ and $g'(x)$ exist for all $x \in [a, b]$, and if $g'(x)$ does not vanish anywhere on $]a, b[$. Then prove that for some c between a and b,

$$\frac{f(c) - f(a)}{g(b) - g(c)} = \frac{f'(c)}{g'(c)}$$

[*Hint*: Apply Rolle's Theorem to the function $fg - gf(a) - fg(b)$].

12. Apply Lagrange's Mean Value Theorem to the function $\log(1 + x)$ to show that

$$0 < [\log(1+x)]^{-1} - x^{-1} < 1, \ \forall x > 0$$

13. If $\left| f(x) - f(y) \right| \leq (x-y)^2$, for all real numbers x and y. Prove that f is a constant function.

14. Applying Lagrange's Mean Value Theorem to the function

$$f(x) = \tan^{-1} x,\ x \in \mathbf{R}.$$

Show that f and f' both are uniformly continuous on $\mathbf{R}$.

15. Find the range of the values of x for which the function $x^3 - 6x^2 - 36x + 7$, increases with x.

16. Establish the following inequalities :

(i) $x - \dfrac{x^2}{2} + \dfrac{x^3}{3(1+x)} < \log(1+x) < x - \dfrac{x^2}{2} + \dfrac{x^3}{3},\ x > 0$

(ii) $\dfrac{x^2}{2(1+x)} < x - \log(1+x) < \dfrac{x^2}{2},\ x > 0$

(iii) $\dfrac{x^2}{2} < x - \log(1+x) < \dfrac{x^2}{2(1+x)},\ -1 < x < 0$

(iv) $1 - x < -\log x < \dfrac{1}{x} - 1,\ 0 < x < 1$

(v) $x < -\log(1-x) < x(1-x)^{-1},\ 0 < x < 1$

(vi) $x < \sin^{-1} x < \dfrac{x}{\sqrt{1-x^2}},\ 0 < x < 1$

(vii) $(1-x) < e^{-x} < 1 - x + \dfrac{x^2}{2},\ x > 0$

(viii) $1 + x \leq e^x \leq 1 + xe^x,\ \forall\, x$

(ix) $x/(1+x) \leq \log(1+x) \leq x,\ x > -1$

17. If $g(x) = 0$ has two equal roots, show that $g'(x) = 0$ has one root equal to either.

ANSWERS

1. (i), (ii) Both valid, (iii), (iv) and (v) not valid.

2. (i) Not valid, (ii) valid, (iii) valid, $c = (6 - \sqrt{21})/6$,

(iv) Hypothesis not valid but the conclusion is valid,

(v) valid, $c = \frac{7}{2}$.

15. $x < -2$ and $x > 6$.

8. HIGHER ORDER DERIVATIVES

We know that the existence of the derivative f' of a function f at a point c implies the existence and continuity of the function in a neighbourhood of c. The derivative of the function f' at c in case it

exists, is called the second derivative of f at c and denoted by $f''(c)$. Evidently the existence of $f''(c)$ implies the existence and continuity of f' in a neighbourhood of c.

Higher order derivatives can be similarly defined. The derivative of f^{n-1} at c, in case, it exists is called the nth derivative of f at c and is denoted by $f^n(c)$.

8.1 Taylor's Theorem

If a function f defined on $[a, a+h]$, is such that (i) the $(n-1)$th derivative f^{n-1} is continuous on $[a, a+h]$, and (ii) the nth derivative f^n exists on $]a, a+h[$, then $\exists$ at least one real number θ between 0 and 1 $(0 < \theta < 1)$ such that

$$f(a+h) = f(a) + hf'(a) + \frac{h^2}{2!}f''(a) + \frac{h^3}{3!}f'''(a) + \ldots$$

$$\ldots + \frac{h^{n-1}}{(n-1)!}f^{n-1}(a) + \frac{h^n(1-\theta)^{n-p}}{p[(n-1)!]}f^n(a+\theta h) \qquad \ldots(1)$$

where p is a given positive integer.

First of all we observe that the condition (i) in the statement implies that all the derivatives $f', f'', \ldots, f^{n-1}$ exist and are continuous on $[a, a+h]$.

Consider the function ϕ defined on $[a, a+h]$ as

$$\phi(x) = f(x) + (a+h-x)f'(x) + \frac{(a+h-x)^2}{2!}f''(x) + \ldots$$

$$\ldots + \frac{(a+h-x)^{n-1}}{(n-1)!}f^{n-1}(x) + A(a+h-x)^p$$

where A is a constant to be determined such that $\phi(a+h) = \phi(a)$.

$$\therefore \qquad f(a+h) = f(a) + hf'(a) + \frac{h^2}{2!}f''(a) + \ldots$$

$$\ldots + \frac{h^{n-1}}{(n-1)!}f^{n-1}(a) + Ah^p \qquad \ldots(2)$$

Now,

(i) $f, f', f'', \ldots, f^{n-1}$ being all continuous on $[a, a+h]$, the function $\phi(x)$ is continuous on $[a, a+h]$;

(ii) the functions $f, f', \ldots, f^{n-1}$ and $(a+h-x)^r$ for all r being derivable in $]a, a+h[$, the function $\phi(x)$ is derivable in $]a, a+h[$; and

(iii) $\phi(a+h) = \phi(a)$.

Thus, the function $\phi(x)$ satisfies all the conditions of Rolle's Theorem and hence $\exists$ at least one real number θ between 0 and 1 such that $\phi'(a+\theta h) = 0$.

But

$$\phi'(x) = \frac{(a+h-x)^{n-1}}{(n-1)!} f^n(x) - Ap(a+h-x)^{p-1}$$

$$\therefore \qquad 0 = \phi'(a+\theta h) = \frac{h^{n-1}(1-\theta)^{n-1}}{(n-1)!} f^n(a+\theta h) - Aph^{p-1}(1-\theta)^{p-1}$$

$$\Rightarrow \qquad A = \frac{h^{n-p}(1-\theta)^{n-p}}{p(n-1)!} f^n(a+\theta n), \; h \neq 0, \theta \neq 1 \qquad ...(3)$$

Substituting A from (3) in (2), we get the required result.

Forms of Remainder after n Terms

(*i*) The term

$$R_n = \frac{h^n(1-\theta)^{n-p}}{p[(n-1)!]} f^n(a+\theta h)$$

which occurs after n terms, is known as Taylor's remainder after n terms. The theorem with this form of remainder is known as Taylor's Theorem with *Schlömilch and Röche form of remainder*.

(*ii*) For $p = 1$, we get

$$R_n = \frac{h^n(1-\theta)^{n-1}}{(n-1)!} f^n(a+\theta h)$$

called *Cauchy's form of remainder*.

(*iii*) For $p = n$, we get

$$R_n = \frac{h^n}{n!} f^n(a+\theta h)$$

called *Lagrange's form of remainder*.

EXERCISE

1. Prove Taylor's Theorem with Lagrange's form of remainder by considering the function

$$\phi(x) = f(x) + (a+h-x) f'(x) + \frac{(a+h-x)^2}{2!} f''(x) + ...$$

$$...+ \frac{(a+h-x)^{n-1}}{(n-1)!} f^{n-1}(x) + A(a+h-x)^n$$

2. Prove Taylor's Theorem with Cauchy's form of remainder by taking the last term of $\phi(x)$ as $A(a+h-x)$.

Second form of Taylor's Theorem. If f satisfies the conditions of Taylor's Theorem in $[a, a+h]$ and x is any point of $[a, a+h]$ then it satisfies the conditions in the interval $[a, x]$ also.

Replacing $a+h$ by x or h by $(x-a)$ in (1), we get

$$f(x) = f(a) + (x-a)f'(a) + \frac{(x-a)^2}{2!}f''(a) + \ldots + \frac{(x-a)^{n-1}}{(n-1)!}f^{n-1}(a)$$

$$+\frac{(x-a)^n}{p[n-1]!}(1-\theta)^{n-p} f^n(a+\theta(x-a)) \qquad \ldots(4)$$

where $0 < \theta < 1$.

The remainder after n terms can thus be written as

$$R_n = \frac{(x-a)^n(1-\theta)^{n-p}}{p[(n-1)!]} f^n(c)$$

where c lies between a and x and depends on the selection of x.

8.2 Maclaurin's Theorem

Putting $a = 0$ in (4), we have for $x \in]0, h[$

$$f(x) = f(0) + xf'(0) + \frac{x^2}{2!}f''(0) + \ldots + \frac{x^{n-1}}{(n-1)!}f^{n-1}(0)$$

$$\frac{x^n(1-\theta)^{n-p}}{p[(n-1)!]} f^n(\theta x)$$

is called Maclaurin's Theorem with *Schlömilch* and *Röche* form of remainder.

Cauchy's form of remainder (for $p = 1$):

$$R_n = \frac{x^n(1-\theta)^{n-1}}{(n-1)!} f^n(\theta x)$$

Lagrange's form of remainder (for $p = n$):

$$R^n = \frac{x^n}{n!} f^n(\theta x)$$

We have thus proved *Maclaurin's Theorem*. Thus Maclaurin's Theorem with Lagrange's form of remainder may be stated as:

If f^{n-1} is continuous in $[0, h]$ *and is derivable in* $]0, h[$, *then for each* $x \in [0, h]$, *there exists a number* θ *between* 0 *and* 1 *such that*

$$f(x) = f(0) + xf'(0) + \frac{x^2}{2!}f''(0) + \ldots + \frac{x^{n-1}}{(n-1)!}f^{n-1}(0) + \frac{x^n}{n!}f^n(\theta x).$$

8.3 Generalised Mean Value Theorem (*Taylor's theorem*)

Deduction of Taylor's Theorem from the Mean Value Theorem

Let a function f be such that $(n-1)$th derivative f^{n-1} is continuous in $[a, a+h]$ and its nth derivative f^n exists in $]a, a+h[$. Consequently, the functions $f, f', f'' \ldots f^{n-1}$ exist and are continuous in $[a, a+h]$ while f^n exists in $]a, a+h[$.

Consider the function

$$\phi(x) = f(x) + (a+h-x)f'(x) + \frac{(a+h-x)^2}{2!}f''(x) + \ldots$$

$$\ldots + \frac{(a+h-x)^{n-1}}{(n-1)!}f^{n-1}(x)$$

which, being the sum of continuous and derivable functions, is itself continuous in $[a, a+h]$ and derivable in $]a, a+h[$. Therefore by Lagrange's Mean Value Theorem $\exists$ a positive number θ between 0 and 1 such that

$$\phi(a+h) = \phi(a) + h\phi'(a+\theta h)$$

Now

$$\phi'(x) = \frac{(a+h-x)^{n-1}}{(n-1)!}f^n(x)$$

$$\therefore \qquad \phi'(a+\theta h) = \frac{h^{n-1}(1-\theta)^{n-1}}{(n-1)!}f^n(a+\theta h)$$

Also

$$\phi(a) = f(a) + hf'(a) + \frac{h^2}{2!}f''(a) + \ldots + \frac{h^{n-1}}{(n-1)!}f^{n-1}(a)$$

and

$$\phi(a+h) = f(a+h)$$

$$\therefore \qquad f(a+h) = f(a) + hf'(a) + \frac{h^2}{2!}f''(a) + \ldots + \frac{h^{n-1}}{(n-1)!}f^{n-1}(a) + \frac{h^n(1-\theta)^{n-1}}{(n-1)!}f^n(a+\theta h)$$

where $0 < \theta < 1$, which is *Taylor's Theorem with Cauchy's form of remainder*.

Note: For $n = 1$, the theorem reduces to the Mean Value Theorem. For this reason Taylor's Theorem is also called General Mean Value Theorem.

8.4 Taylor's Infinite Series and Power Series Expansions

We have seen that

$$f(a+h) = f(a) + hf'(a) + \frac{h^2}{2!}f''(a) + \ldots + \frac{h^{n-1}}{(n-1)!}f^{n-1}(a) + R_n \qquad \ldots(5)$$

where R_n is the remainder after n terms.

The result can be interpreted in two ways:

(*i*) The value $f(a + h)$ of the function at a point may be approximated by a summation of the terms like $\frac{h^r}{r!} f^r(a)$ involving values of the function and its derivatives at some other point of the domain of definition.

(*ii*) The value $f(a + h)$ of the function may be expanded in powers of h.

The natural question as to how far the R.H.S. of (5) correctly represents the L.H.S. is answered if the series

$$f(a) + hf'(a) + \frac{h^2}{2!} f''(a) + \ldots$$

converges to $f(a + h)$.

Let $S_n = f(a) + hf'(a) + \frac{h^2}{2!} f''(a) + \ldots + \frac{h^{n-1}}{(n-1)!} f^{n-1}(a)$, so that

$$f(a + h) = S_n + R_n$$

Thus if $R_n \to 0$ as $n \to \infty$, we have

$$\lim_{n\to\infty} S_n = f(a + h)$$

i.e., the infinite series

$$f(a) + hf'(a) + \frac{h^2}{2!} f''(a) + \ldots + \frac{h^n}{n!} f^n(a) + \ldots$$

converges to $f(a + h)$.

Thus we have proved that if a function f possesses derivatives of every order in $[a, a + h]$ and Taylor's remainder $R_n \to 0$ as $n \to \infty$, then

$$f(a) + hf'(a) + \frac{h^2}{2!} f''(a) + \ldots = f(a + h)$$

The infinite series

$$f(a) + hf'(a) + \frac{h^2}{2!} f''(a) + \frac{h^3}{3!} f'''(a) + \ldots \qquad \ldots(6)$$

is called *Taylor's series*. It can also be looked upon as expansion of $f(a + h)$ in powers of h.

Similarly for $x \in [a, a + h]$, when $\lim_{n\to\infty} R_n = 0$, we have from (4),

$$f(x) = f(a) + (x - a) f'(a) + \frac{(x-a)^2}{2!} f''(a) + \ldots \qquad \ldots(7)$$

which is the expansion of $f(x)$ in powers of $(x - a)$.

8.5 Maclaurin's Infinite Series

We may easily deduce from (5) or (6) that if f possesses derivatives of every order in $[0, h[$ and $\lim_{n\to\infty} R_n = 0$, then for all $x \in [0, h]$,

$$f(x) = f(0) + xf'(0) + \frac{x^2}{2!} f''(0) + \ldots$$

which is *Maclaurin's infinite series* expansion of $f(x)$ in powers of x.

Note: In the above discussion the remainder R_n can be in any of the forms.

Example 12. Show that the number θ which occurs in the Taylor's Theorem with Lagrange's form of remainder after n terms approaches the limit $1/(n+1)$ as h approaches zero, provided that $f^{n+1}(x)$ is continuous and different from zero at $x = a$.

■ Applying Taylor's Theorem with remainders after n terms and $n + 1$ terms successively, we get

$$f(a+h) = f(a) + hf'(a) + \ldots + \frac{h^{n-1}}{(n-1)!} f^{n-1}(a) + \frac{h^n}{n!} f^n(a + \theta h)$$

$$f(a+h) = f(a) + hf'(a) + \ldots + \frac{h^n}{n!} f^n(a) + \frac{h^{n+1}}{(n+1)!} f^{n+1}(a + \theta' h)$$

These give

$$\frac{h^n}{n!} f^n(a + \theta h) = \frac{h^n}{n!} f^n(a) + \frac{h^{n+1}}{(n+1)!} f^{n+1}(a + \theta' h)$$

or

$$f^n(a + \theta h) - f^n(a) = \frac{h}{n+1} f^{n+1}(a + \theta' h)$$

Applying Lagrange's Mean Value Theorem to the left hand side, we have

$$\theta h f^{n+1}(a + \theta'' \theta h) = \frac{h}{n+1} f^{n+1}(a + \theta' h)$$

or

$$\theta = \frac{1}{n+1} \frac{f^{n+1}(a + \theta' h)}{f^{n+1}(a + \theta'' \theta h)}$$

Taking the limit when $h \to 0$, we get

$$\lim_{h \to 0} \theta = \frac{1}{n+1}$$

8.6 To illustrate the applications of these theorems we consider series *expansion of the functions* e^x, $\cos x$, $\log(1 + x)$, $(1 + x)^m$ by Maclaurin's Theorem.

1. Let $f(x) = e^x$ so that $f^n(x) = e^x, \ \forall\, n$.

   Evidently $f(x)$ and all its derivatives exist and are continuous for every real value of x.

   Let us now consider the limit of the remainder R_n.

   Taking Lagrange's form of the remainder, we have

$$R_n = \frac{x^n}{n!} f^n(\theta x) = \frac{x^n}{n!} e^{\theta x}, 0 < \theta < 1$$

$$\therefore \qquad \lim_{n\to\infty} R_n = \lim_{n\to\infty} \frac{x^n}{n!} e^{\theta x} = \left(\lim_{n\to\infty} \frac{x^n}{n!}\right) e^{\theta x} = 0$$

Thus the conditions of Maclaurin's infinite expansion are satisfied. Now $f(0) = 1$ and $f^n(0) = 1$ for all integral values of n.

Substituting these values in the Maclaurin's infinite series, we have

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \ldots \quad \forall\ x \in \mathbf{R}$$

2. Let $f(x) = \cos x$, so that $f^n(x) = \cos\left(\frac{1}{2}n\pi + x\right), \forall\ n.$

Evidently $f(x)$ and all its derivatives exist and are continuous for every real value of x. Taking Lagrange's form of the remainder,

$$R_n = \frac{x^n f^n(\theta x)}{n!} = \frac{x^n}{n!} \cos\left(\frac{n\pi}{2} + \theta x\right)$$

$$\Rightarrow \qquad |R_n| = \left|\frac{x^n}{n!}\right| . \left|\cos\left(\frac{n\pi}{2} + \theta x\right)\right| \le \left|\frac{x^n}{n!}\right|$$

$$\therefore \qquad \lim_{n\to\infty} R_n = 0, \ \forall\ x$$

Thus the conditions of Maclaurin's infinite expansion are satisfied.

Now

$$f(0) = 1,\ f'(0) = 0,\ f''(0) = -1, \ldots,\ f^n(0) = \cos(n\pi/2)$$

which is zero for n odd, and alternately $+1, -1$ for n even.

Substituting in the equation

$$f(x) = f(0) + x\, f'(0) + \frac{x^2}{2!} f''(0) + \ldots$$

we get

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \ldots \quad \forall x \in R$$

3. Let $f(x) = \log(1 + x)$, for $-1 < x \le 1$, then

$$f^n(x) = \frac{(-1)^{n-1}(n-1)!}{(1+x)^n}$$

Evidently $f(x)$ and all its derivatives exist and are continuous for $|x| < 1$.

Taking the Lagrange's form of remainder, we have

$$R_n = \frac{x^n}{n!} f^n(\theta x) = \frac{(-1)^{n-1} x^n}{n(1+\theta x)^n} = (-1)^{n-1} \frac{1}{n}\left(\frac{x}{1+\theta x}\right)^n$$

(a) When $0 \le x \le 1$, then $0 < \theta x < x \le 1$ and

$$|R_n| = \frac{x^n}{n}\left(\frac{1}{1+\theta x}\right)^n \le \frac{x^n}{n} \le \frac{1}{n}.$$

Also $\frac{1}{n} \to 0$ as $n \to \infty$

$\therefore$ $R_n \to 0$ as $n \to \infty$, for $0 \le x \le 1$.

Thus, the conditions of Maclaurin's infinite series expansion are satisfied for $0 \le x \le 1$.

(b) When $-1 < x < 0$.

In this case x may or may not be numerically less than $1 + \theta x$, so that nothing can be said about the limit of $\left(\frac{x}{1+\theta x}\right)^n$ when $n \to \infty$. Thus we fail to draw any definite conclusion from Lagrange's form of remainder. Let us now see if the other form of the remainder is of any help.

Cauchy's form of remainder,

$$R_n = \frac{x^n (1-\theta)^{n-1} f^n(\theta x)}{(n-1)!} = \frac{(-1)^{n-1} x^n (1-\theta)^{n-1}}{(1+\theta x)^n}$$

$$= (-1)^{n-1} x^n \left(\frac{1-\theta}{1+\theta x}\right)^{n-1} \cdot \frac{1}{(1+\theta x)}$$

Now $(1-\theta) < 1 + \theta x$ so that $\left(\frac{1-\theta}{1+\theta x}\right)^{n-1} \to 0$ as $n \to \infty$.

Also

$$x^n \to 0 \text{ as } n \to \infty \text{ and } \frac{1}{1+\theta x} < \frac{1}{1-|x|}$$

and moreover it is independent of n.

Thus, $R_n \to 0$ as $n \to \infty$.

Hence, the conditions of Maclaurin's series expansion are satisfied also when $-1 < x < 0$.

Thus substituting the values $f(0) = 0$, $f^n(0) = (-1)^{n-1}[(n-1)!]$ in Maclaurin's infinite expansion, we get

$$\log(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \dots, \text{ for } -1 < x \le 1$$

4. Let $f(x) = (1+x)^m$.

Two cases arise according as m is or not a positive integer.

(a) When m is a positive integer, $f(x)$ possesses continuous derivatives of all orders upto m. Derivatives of order higher than m vanish identically. Consequently for $n > m$, $R_n \to 0$ identically so that the conditions of Maclaurin's expansion are satisfied. On substituting, the

values of $f(0), f'(0), f''(0),..., f^m(0)$, we get the finite expansion

$$(1+x)^m = 1 + mx + \frac{m(m-1)}{2!}x^2 + ... + x^m, \ \forall\, x$$

(*b*) When m is not a positive integer, $(1+x)^m$ possesses continuous derivatives of all orders provided $x \neq -1$.

Let $-1 < x < 1$.

Taking Cauchy's form of remainder, we have

$$\begin{aligned} R_n &= \frac{x^n}{(n-1)!}(1-\theta)^{n-1} f^n(\theta x) \\ &= \frac{x^n}{(n-1)!}(1-\theta)^{n-1} m(m-1)\ldots(m-n+1)\,(1+\theta x)^{m-n} \\ &= \left(\frac{m(m-1)\ldots(m-n+1)\,x^n}{(n-1)!}\right)\left(\frac{1-\theta}{1+\theta x}\right)^{n-1}(1+\theta x)^{m-1} \end{aligned}$$

We know for $|x| < 1$,

$$\frac{m(m-1)(m-2)\ldots(m-n+1)}{(n-1)!}x^n \to 0 \text{ as } n \to \infty,$$

and

$$\frac{1-\theta}{1+\theta x} < 1, \text{ so that } \left(\frac{1-\theta}{1+\theta x}\right)^{n-1} \to 0 \text{ as } n \to \infty.$$

Also

$$(1+\theta x)^{m-1} < (1+|x|)^{m-1}, \ m > 1, \ 0 < \theta < 1$$

and

$$(1+\theta x)^{m-1} = \frac{1}{(1+\theta x)^{1-m}} < \frac{1}{(1-|x|)^{1-m}}, \text{ when } m < 1$$

Thus $R_n \to 0$ when $n \to \infty$, for $|x| < 1$.

Hence, the conditions of Maclaurin's infinite expansion are satisfied.

Making the substitutions $f(0) = 1, f'(0) = m, ..., f^n(0) = m(m-1)\ldots(m-n+1)$, we get

$$(1+x)^m = 1 + mx + \frac{m(m-1)}{2!}x^2 + \frac{m(m-1)(m-2)}{3!}x^3 + \ldots \text{ for } |x| < 1$$

Note: When m is not a positive integer the expansion is not possible if $|x| > 1$, for then as

$$n \to \infty, \frac{m(m-1)\ldots(m-n+1)\,x^n}{(n-1)!} \text{ and so } R_n \text{ does not tend to zero.}$$

EXERCISE

1. Expand, if possible, $\sin x$ in ascending powers of x.

2. Assuming the validity of expansion, show that

   (i) $e^x \cos x = 1 + x - \dfrac{2x^3}{3!} - \dfrac{2^2 x^4}{4!} - \dfrac{2^2 x^5}{5!} + \ldots$

   (ii) $\log \sec x = \dfrac{1}{2}x^2 + \dfrac{1}{12}x^4 + \ldots$

   (iii) $\tan^{-1} x = \tan^{-1}\dfrac{\pi}{4} + \dfrac{x - \pi/4}{1 + \pi^2/16} - \dfrac{\pi(x - \pi/4)^2}{4(1 + \pi^2/16)^2} + \ldots$

   (iv) $\sin\left(\dfrac{\pi}{4} + \theta\right) = \dfrac{1}{\sqrt{2}}\left(1 + \theta - \dfrac{\theta^2}{2!} - \dfrac{\theta^3}{3!} + \ldots\right)$

   (v) $$f(x) = f(a) + 2\left[\frac{x-a}{2} f'\left(\frac{x+a}{2}\right) + \frac{(x-a)^3}{8.(3)!} f'''\left(\frac{x+a}{2}\right) + \frac{(x-a)^5}{32(5)!} f^{\mathrm{v}}\left(\frac{x+a}{2}\right) + \ldots\right]$$

3. Use Taylor's theorem to show that

   (i) $\cos x \geq 1 - \dfrac{x^2}{2}$, for all real x.

   (ii) $x - \dfrac{x^3}{6} < \sin x < x$, for $x > 0$

   (iii) $x - \dfrac{x^3}{6} < \sin x < x - \dfrac{x^3}{6} + \dfrac{x^5}{120}$, $\forall\, x > 0$

   (iv) $1 + x + \dfrac{x^2}{2} < e^x < 1 + x + \dfrac{x^2}{2} e^x$, $x > 0$

4. If $0 < x \leq 2$, then prove that

$$\log x = (x-1) - \frac{(x-1)^2}{2} + \frac{(x-1)^3}{3} - \frac{(x-1)^4}{4} + \ldots$$

5. If $f(x) = \exp(-1/x^2)$, for $x \neq 0$ and $f(0) = 0$, then show that

   (i) $f^n(0) = 0$, for all $n = 0, 1, 2, \ldots$, and

   (ii) The Taylor's series for f about 0 agrees with $f(x)$ only at $x = 0$.

   [*Hint*: First, prove by induction that, for any $x \neq 0$,

$$f^n(x) = \exp(-1/x^2)\, P_n(1/x),$$

   where P_n is a polynomial of degree $3n$. Second, using $e^x > x^n/n!\ (x > 0)$, show that

$$\lim_{x \to 0} \exp(-1/x^2)\, P(1/x) = 0,$$

   where P is any polynomial. Then apply induction to prove (*i*).]

7 Applications of Taylor's Theorem

MAXIMA AND MINIMA, INDETERMINATE FORMS

In this chapter we shall discuss applications of Taylor's theorem to two types of problems: (*a*) extreme values of a function, and (*b*) the evaluation of certain limits, popularly known as Indeterminate Forms.

1. EXTREME VALUES (*Definitions*)

Let c be an interior point of the domain $[a, b]$ of a function f.

Definition 1. The point c is said to be the *stationary point* and $f(c)$, the *stationary value* of the function f if $f'(c) = 0$.

Definition 2. The function f is said to have a *maximum value* (a *maxima* or a *maximum*) at c if $f(c)$ is the greatest value of the function in a small neighbourhood $]c - \delta, c + \delta[$, $\delta > 0$ of c.

Thus for all $x \in]c - \delta, c + \delta]$, $x \neq c$, we have

$$f(c) > f(x)$$

$\Rightarrow$ $f(x) - f(c)$ is negative, for all values of x in $]c - \delta, c + \delta[$ other than c.

Definition 3. $f(c)$ is said to be the *minimum value*, a *minimum* or a *minima* if $f(c)$ is the least value of the function in a small neighbourhood $]c - \delta, c + \delta[$ of c. Thus for all $x \in]c - \delta, c + \delta[$, $x \neq c$,

$$f(c) < f(x)$$

$\Rightarrow$ $f(x) - f(c)$ is positive, for all values of x in a deleted neighbourhood $]c - \delta, c + \delta[$ of c.

Definition 4. The function f is said to have an *extreme value* at c if $f(c)$ is either a maximum or a minimum value. Thus at an extreme point c, $f(x) - f(c)$ keeps the same sign, for all values of x in a deleted neighbourhood $]c - \delta, c + \delta[$ of c.

1.1 A Necessary Condition for Extreme Values

Theorem 1. *If $f(c)$ is an extreme value of a function f then $f'(c)$, in case it exists, is zero.*

For the sake of definiteness, let us assume that $f(c)$ is a maximum value.

Hence $\exists$ a $\delta > 0$ such that

$$f(x) < f(c), \ \forall\, x \in]c - \delta, c + \delta[, \ x \neq c$$

In case $f'(c)$ exists, there are three possibilities:

$$f'(c) > 0, \ f'(c) < 0, \ f'(c) = 0$$

If $f'(c) > 0$, then $\exists$ an interval $]c, c + \delta_1[$, $0 < \delta_1 < \delta$ for every point of which $f(x) > f(c)$, which contradicts the fact that $f(c)$ is a maximum value.

Again if $f'(c) < 0$, then $\exists$ an interval $]c - \delta_2, c[$, $0 < \delta_2 < \delta$ for every point of which $f(x) > f(c)$ which again is a contradiction.

Hence, the only possibility, $f'(c) = 0$.

Remarks:

1. The above theorem states that if the derivative exists, it must vanish at the extreme value. A function may however have an extreme value at a point without being derivable there at. Consider, for example, the function $f(x) = |x|$, which has a minimum at the origin even though $f'(0)$ does not exist.
2. The vanishing of the derivative at a point is only a necessary condition for the existence of an extreme value, it is not sufficient. Functions exist for which the derivative vanishes at a point but do not have an extreme value there at, *e.g.*, $f(x) = x^3$ at $x = 0$, so that the stationary points are not necessarily points of extreme values.

1.2 Investigation of the Points of Maximum and Minimum Values

At a point of extreme value the derivative of the function either does not exist or in case it exists, it must vanish.

Let c be an extreme point of a function f with domain $[a, b]$. If c is a *point of maximum value*, then $\exists$ a neighbourhood $]c - \delta, c + \delta[$ of c such that

$$f(x) < f(c), \quad \forall\, x \in \,]c - \delta, c + \delta[, \; x \neq c$$

$$\therefore \qquad f(x) < f(c), \quad \forall\, x \in \,]c - \delta, c[$$

$\Rightarrow$ f is increasing (to $f(c)$) in a small interval to the left of c.

Again

$$f(x) < f(c), \quad \forall\, x \in \,]c, c + \delta[$$

$\Rightarrow$ f is decreasing (from $f(c)$) in a small interval to the right of c.

We may therefore state:

c is a point of maximum value if the function changes from an increasing to a decreasing function as x passes through c. Therefore, in case f is derivable, the derivative changes sign from positive to negative as x passes through c.

It can similarly be seen that c is a point of *minimum value* if f changes from a decreasing to an increasing function and in case f is derivable, the derivative f' changes sign from negative to positive as x passes through c.

Notes:

1. It appears that f' is a decreasing (increasing) function in a neighbourhood of the point of maxima (minima) and therefore the second derivative f'' in case it exists, would be negative (positive) at such a point.
2. The above conditions are sufficient but not necessary.

1.3 Before discussing the subject further, let us investigate the extreme values in a few cases.

Example 1. Examine the function $(x - 3)^5 (x + 1)^4$ for extreme values.

■ Let $f(x) = (x - 3)^5 (x + 1)^4$.

$$\therefore \qquad f'(x) = (x - 3)^4 (x + 1)^3 (9x - 7)$$

The function is derivable for all $x \in \mathbf{R}$ and the derivative f' vanishes for $x = -1, 3, \frac{7}{9}$ which may now be tested for extreme values.

(*a*) $x = -1$

f' is positive for a value of x slightly less than -1, and negative for slightly greater than -1. Thus f' changes sign from + to – as x passes through -1.

Hence, -1 is a point of maximum value.

(*b*) $x = 3$

f' remains positive as x passes through 3.

Hence, $x = 3$ is neither a maxima nor a minima.

(*c*) $x = \frac{7}{9}$

Since f' changes from – to + as x passes through $\frac{7}{9}$, therefore, f has a minimum value at $x = \frac{7}{9}$.

Example 2. Prove that a conical tent of a given capacity will require the least amount of canvas when the height is $\sqrt{2}$ times the radius of the base.

■ Let the tent be a cone of semi-vertical angle α and radius of the base r. The volume V is fixed, while the surface area S has to be the minimum.

Now

$$V = \frac{1}{3}\pi r^3 \cot \alpha \qquad \text{...(1)}$$

and

$$S = \pi r^2 \operatorname{cosec} \alpha \qquad \text{...(2)}$$

Differentiating (1), we get

$$\frac{\pi}{3}\left[3r^2 \cot \alpha \frac{dr}{d\alpha} - r^3 \operatorname{cosec}^2 \alpha\right] = 0$$

or

$$\frac{dr}{d\alpha} = \frac{r \operatorname{cosec}^2 \alpha}{3 \cot \alpha}$$

Again from (2),

$$\frac{dS}{d\alpha} = \pi\left[2r \operatorname{cosec} \alpha \frac{dr}{d\alpha} - r^2 \cot \alpha \operatorname{cosec} \alpha\right]$$

$$= \frac{\pi r^2 \operatorname{cosec} \alpha}{3 \cot \alpha}\left[2 \operatorname{cosec}^2\alpha - 3 \cot^2\alpha\right]$$

$$= \frac{\pi r^2}{3 \cos \alpha}\left[2 - \cot^2\alpha\right]$$

$$\therefore \qquad \frac{dS}{d\alpha} = 0, \text{ when } \cot\alpha = \sqrt{2}$$

$$\Rightarrow \qquad \alpha = \cot^{-1}\sqrt{2}.$$

Also $dS/d\alpha$ changes sign from negative to positive as α passes through the value $\cot^{-1}\sqrt{2}$.

Therefore S has a minimum value at $\alpha = \cot^{-1}\sqrt{2}$.

Hence, the height of the tent $= r\cot\alpha = r\sqrt{2} = \sqrt{2}$ times the radius of the base.

1.4 In most cases the conclusion of § 1.2 suffices to determine the points of extreme values but to simplify the matters, recourse may be had to higher derivatives. In that context the following theorem will prove very useful.

Theorem 2. *If c is an interior point of the domain of a function f and $f'(c) = 0$, then the function has a maxima or a minima at c according as $f''(c)$ is negative or positive.*

The existence of $f''(c)$ implies that f and f' exist and are continuous at c. Continuity at c implies the existence of f and f' in a certain neighbourhood $]c - \delta, c + \delta[, \delta > 0$ of c, the neighbourhood being itself contained in the domain of f.

Let $f''(c) > 0$.

This implies that $f'(x)$ is an increasing function of c.

$$\Rightarrow \qquad f'(x) > f'(c) = 0, \;\; \forall\, x \in \left]c, c+\delta_1\right[, \delta_1 < \delta$$

and

$$f'(x) < f'(c) = 0 \;\; \forall\, x \in \left]c - \delta_1, c\right[$$

Thus first implies that $f'(x)$ is positive and hence f is an increasing function in $]c, c + \delta_1[$, *i.e.*, $f(x) > f(c)$ in $]c, c + \delta_1[$.

Similarly $f(x) > f(c)$ in $]c - \delta_1, c[$.

The last two results imply that

$$f(x) > f(c), \forall\, x \in \left]c - \delta_1, c + \delta_1\right[, x \neq c.$$

$\Rightarrow \qquad f$ has a minima at c.

Similarly $f''(c) < 0 \Rightarrow f(x)$ has a maxima at c.

Example 3. Examine the function $\sin x + \cos x$ for extreme values.

■ Let

$$f(x) = \sin x + \cos x$$
$$f'(x) = \cos x - \sin x$$
$$f''(x) = -\sin x - \cos x$$
$$f'(x) = 0 \text{ when } \tan x = 1, \text{ so that}$$
$$x = n\pi + \frac{\pi}{4}$$

where n is zero or any integer

$$f''\left(n\pi + \frac{1}{4}\pi\right) = -\left\{\sin\left(n\pi + \frac{1}{4}\pi\right) + \cos\left(n\pi + \frac{1}{4}\pi\right)\right\}$$

$$= (-1)^{n+1}\left(\sin\frac{\pi}{4} + \cos\frac{\pi}{4}\right) = (-1)^{n+1}\sqrt{2}$$

Also $$f\left(n\pi + \frac{1}{4}\pi\right) = \sin\left(n\pi + \frac{1}{4}\pi\right) + \cos\left(n\pi + \frac{1}{4}\pi\right) = (-1)^n\sqrt{2}$$

When n is zero or an even integer, $f''(n\pi + \pi/4)$ is negative and therefore $x = n\pi + \frac{1}{4}\pi$ makes $f(x)$ a maxima with the maximum value $\sqrt{2}$.

When n is an odd integer, $f''(n\pi + \frac{1}{4}\pi)$ is positive and therefore $x = n\pi + \frac{1}{4}\pi$ makes $f(x)$ a minima with the minimum value $\sqrt{2}$.

1.5* In cases where the second derivative vanishes, the above theorem fails to give any result. In those cases, we make use of still higher derivatives, and the following theorem proves very helpful.

General Theorem. *If c is an interior point of the domain $[a, b]$ of a function f and is such that*

(i) $f'(c) = f''(c) = f'''(c) = ... = f^{n-1}(c) = 0$, *and*

(ii) $f^n(c)$ *exists and is not zero*,

then for n odd, $f(c)$ is not an extreme value, while for n even $f(c)$ is a maximum or minimum value according as $f^n(c)$ is negative or positive.

Condition (*ii*) of the existence of $f^n(c)$ implies that $f, f', f'', ..., f^{n-1}$ all exist and are continuous at c. Also continuity at c implies the existence of $f, f', ..., f^{n-1}$ in a certain neighbourhood $]c - \delta_1, c + \delta_1[$ of c, $\delta_1 > 0$.

As $f^n(c) \neq 0$ there exists a neighbourhood $]c - \delta, c + \delta[$, $0 < \delta < \delta_1$ such that for $f^n(c) > 0$,

and $$\left.\begin{aligned} f^{n-1}(x) < f^{n-1}(c) = 0, \; x \in\,]c - \delta, c[\\ f^{n-1}(x) > f^{n-1}(c) = 0, \; x \in\,]c, c + \delta[\end{aligned}\right\} \quad ...(1)$$

and for $f^n(c) < 0$,

and $$\left.\begin{aligned} f^{n-1}(x) > f^{n-1}(c) = 0, \; x \in\,]c - \delta, c[\\ f^{n-1}(x) < f^{n-1}(c) = 0, \; x \in\,]c, c + \delta[\end{aligned}\right\} \quad ...(2)$$

Again for any real number h, where $|h| < \delta$, we have by Taylor's Theorem

$$f(c + h) - f(c) + hf'(c) + \frac{h^2}{2!}f''(c) + ... + \frac{h^{n-1}}{(n-1)!}f^{n-1}(c + \theta h),\ 0 < \theta < 1$$

or $$f(c + h) - f(c) = \frac{h^{n-1}}{(n-1)!}f^{n-1}(c + \theta h) \quad ...(3)$$

where $c + \theta h \in\,]c - \delta, c + \delta[$.

For n odd: Clearly $h^{n-1} > 0$ for any real number h, and further:

(*i*) When $f^n(c) > 0$, we deduce from (1) that for h negative, $(c + \theta h)$ is in $[c - \delta, c[$, and so $f^{n-1}(c + \theta h) < 0$, and for h positive, $f^{n-1}(c + \theta h) > 0$.

Thus from (3)

$$f(c+h) < f(c), \text{ when } c - \delta < c + h < c$$

and

$$f(c+h) > f(c), \text{ when } c < c + h < c + \delta$$

Thus $f(c)$ is not an extreme value.

(*ii*) When $f^n(c) < 0$, it may similarly be shown that $f(c)$ is not an extreme value.

For n even: As h^{n-1} is positive or negative according as h is positive or negative, we deduce as before from (1) and (3) that if $f^n(c) > 0$ then for every point $x = c + h \in]c - \delta, c + \delta[$ except c,

$$f(c+h) > f(c)$$

i.e., $f(c)$ is a minimum value.

It may similarly be deduced from (1) and (3) that $f(c)$ is a maximum value if $f^n(c) < 0$.

Notes:

1. As a consequence of the above theorem, if f' vanishes at c, then c is a point of maxima if $f''(c) < 0$ and a minima if $f''(c) > 0$.
2. Extreme value exists only if the first non-zero derivative is of even order.

Example 4. Examine the function $(x - 3)^5 (x + 1)^4$ for extreme values.

■ Clearly f is derivable.

Let $f(x) = (x - 3)^5 (x + 1)^4$

$$\therefore \quad f'(x) = (x-3)^4 (x+1)^3 [9x - 7]$$

$$f''(x) = 8(x-3)^3 (x+1)^2 (9x^2 + 14x + 1)$$

$$f'''(x) = 24(x-3)^2 (x+1)(21x^3 - 49x^2 + 7x + 13)$$

$$f^{iv}(x) = 24(x-3)(3x-1)(21x^3 - 49x^2 + 7x + 13) + 168(x+3)^2 (x+1)(9x^2 - 14x + 1)$$

$$f^{v}(x) = 48(3x-5)(21x^3 - 49x^2 + 7x + 13) + 336(x-3)(9x^2 - 14x + 1)(3x-1) + 336(x-3)^2 (x+1)(9x-7)$$

Now f' vanishes for $x = -1, 3, \dfrac{7}{9}$.

Let us now test these for extreme values.

At $x = -1$, f^{iv} is the first non-vanishing derivative and this is negative. Thus $x = -1$ is a point of maxima.

At $x = 3$ the first non-vanishing derivative is the fifth which is of odd order. Therefore, the function has neither maximum nor minimum at $x = 3$. Such a point is called the *point of inflexion* of the function.

At $x = \frac{7}{9}$, f'' is the first non-vanishing derivative and is positive and therefore it is a point of minima.

EXERCISE

1. Show that the maximum value of the function $(x-1)(x-2)(x-3)$ is $\frac{2\sqrt{3}}{9}$ at $x = 2 - \frac{1}{\sqrt{3}}$.
2. Show that $x^5 - 5x^4 + 5x^3 - 1$ has a maxima at $x = 1$ and a minima at $x = 3$ and neither at $x = 0$.
3. Find the maximum and the minimum as well as the greatest and the least value of $x^3 - 12x^2 + 45x$ in the interval $[0, 7]$.
4. Find the maximum or minimum of
$$\frac{x^4}{(x-1)(x-3)^3}.$$
5. Show that the maximum value of $(1/x)^x$ is $e^{1/e}$.
6. Show that the maximum value of $(\log x)/x$ in $0 < x < \infty$ is $1/e$.
7. Show that $\sin x(1 + \cos x)$ is maximum at $x = \pi/3$.
8. If $(x-a)^{2n}(x-b)^{2m+1}$, where m and n are positive integers, is the derivative of a function f then show that $x = b$ gives a minimum but $x = a$ gives neither a maximum nor a minimum.
9. Show that the semi-vertical angle of a cone of maximum volume and of given slant height is $\tan^{-1}\sqrt{2}$.
10. Show that the volume of the greatest cylinder which can be inscribed in a cone of height h and semi-vertical angle α is $(4/27)\pi h^3 \tan^2\alpha$.
11. Show that the height of the cylinder of maximum volume that can be inscribed in a sphere of radius a is $2a/\sqrt{3}$.

ANSWERS

3. Max 54, min 50, greatest 70, least 0.
4. Max at $\frac{6}{5}$, min at $x = 0$.

2. INDETERMINATE FORMS

We shall now discuss the evaluation of limits of functions generally known as *Indeterminate forms*. They are not indeterminate but have acquired this name by usage of the word.

In general, the limit of $\phi(x)/\psi(x)$ when $x \to a$, in case the limits of both the functions exist, is equal to the limit of the numerator divided by the limit of the denominator. But what happens when both these limits are zero? The division (0/0) then becomes meaningless. A case like this is known as Indeterminate form. Other such forms are ∞/∞, $0 \times \infty$, $\infty - \infty$, 0^0, 1^∞, and ∞^0. Ordinary methods of evaluating the limits are of little help. Particular methods are required to evaluate these peculiar limits. We shall now discuss these particular methods, generally called *L' Hospital rule*, due to the French mathematician L' Hospital (also called L' Hospital).

It should, however, be clearly understood, that in what follows, we do not find the value of 0/0 or of any of the other indeterminate forms. We only find the limits of combinations of functions which assume these forms when the limits of functions are taken separately.

2.1 It will be of help to remember the following points concerning the limits and continuity:

(*i*) $\lim_{x \to a} f(x) = l$

i.e., the function f is defined for every point ξ of the deleted neighbourhood $]a - \delta, a + \delta[$ of a and $f(\xi) \to l$ as $\xi \to a$.

(*ii*) Continuity of $f(x)$ at $x = a \Rightarrow \lim_{x \to a} f(x) = f(a)$,

i.e., the function f is defined for every point ξ of a neighbourhood $]a - \delta, a + \delta[$ of a and $f(\xi) \to f(a)$ as $\xi \to a$.

(*iii*) $\lim_{x \to a} \dfrac{f(x)}{g(x)} = l \; (\neq 0 \text{ or } \infty)$, then

(*a*) $\lim_{x \to a} g(x) \neq 0 \; (\text{or } \infty) \Rightarrow \lim_{x \to a} f(x)$ exists finitely

(*b*) $\lim_{x \to a} g(x) = 0 \; (\text{or } \infty) \Rightarrow \lim_{x \to a} f(x) = 0 \; (\text{or } \infty)$

for

$$\lim f(x) = \lim \left(\frac{f(x)}{g(x)} \cdot g(x) \right) = \lim \left(\frac{f(x)}{g(x)} \right) \lim g(x) = l \cdot (\lim g(x))$$

2.2 Indeterminate Form, 0/0

We shall now discuss some theorems concerning the indeterminate form 0/0. The reader will do well to note the differences in the hypothesis and the line of proof of the theorems.

Theorem 3. *If f, g be two functions such that*

(*i*) $\lim_{x \to a} f(x) = \lim_{x \to a} g(x) = 0$ *and*

(*ii*) $f'(a), g'(a)$ *exist and* $g'(a) \neq 0$ *then*

$$\lim_{x \to a} \frac{f(x)}{g(x)} = \frac{f'(a)}{g'(a)}$$

Since the functions f and g are derivable at $x = a$, therefore, they are continuous there at, *i.e.*,

$$\lim_{x \to a} f(x) = f(a) \quad \text{and} \quad \lim_{x \to a} g(x) = g(a)$$

Thus from condition (*i*), $f(a) = 0 = g(a)$.

Also

$$f'(a) = \lim_{x \to a} \frac{f(x) - f(a)}{x - a} = \lim_{x \to a} \frac{f(x)}{x - a}$$

and

$$g'(a) = \lim_{x \to a} \frac{g(x) - g(a)}{x - a} = \lim_{x \to a} \frac{g(x)}{x - a}$$

$\therefore$
$$\frac{f'(a)}{g'(a)} = \lim_{x \to a} \frac{f(x)}{g(x)}$$

Note: Condition (i) can be replaced by $f(a) = g(a) = 0$.

Theorem 4. *L' Hospital's Rule for 0/0 form. If f, g are two functions such that*

(i) $\lim_{x \to a} f(x) = \lim_{x \to a} g(x) = 0,$

(ii) $f'(x), g'(x)$ *exist and* $g'(x) \neq 0, \forall\, x \in \,]a - \delta, a + \delta[\,, \delta > 0$ *except possibly at a, and*

(iii) $\lim_{x \to a} \dfrac{f'(x)}{g'(x)}$ *exists,*

then
$$\lim_{x \to a} \frac{f(x)}{g(x)} = \lim_{x \to a} \frac{f'(x)}{g'(x)}$$

Let us define two functions F and G such that

$$F(x) = \begin{cases} f(x), \ \forall\, x \in \,]a - \delta, a + \delta[\, \text{ except } a \\ \lim\limits_{x \to a} f(x), \text{ at } x = a \end{cases}$$

$$G(x) = \begin{cases} g(x), \ \forall\, x \in \,]a - \delta, a + \delta[\, \text{ except } a \\ \lim\limits_{x \to a} g(x), \text{ at } x = a \end{cases}$$

Clearly F and G are continuous and derivable on $]a - \delta, a + \delta[$ except possibly at a.
Also, since

$$\lim_{x \to a} F(x) = \lim_{x \to a} f(x) = F(a)$$

and

$$\lim_{x \to a} G(x) = \lim_{x \to a} g(x) = G(a)$$

hence F and G are continuous at a as well.

Let x be a point of $]a - \delta, a + \delta[$ other than a.

For $x > a$, F and G satisfy the conditions of Cauchy's Mean Value Theorem in $[a, x]$, so that

$$\frac{F(x)-F(a)}{G(x)-G(a)} = \frac{F'(\xi)}{G'(\xi)}, \text{ where } a < \xi < x$$

But $$F(a) = 0 = G(a)$$

$\therefore$ $$\frac{F(x)}{G(x)} = \frac{F'(\xi)}{G'(\xi)} \qquad ...(1)$$

Proceeding to limits,

$$\lim_{x\to a+0} \frac{F(x)}{G(x)} = \lim_{\xi\to a+0} \frac{F'(\xi)}{G'(\xi)} = \lim_{x\to a+0} \frac{F'(x)}{G'(x)}$$

$\Rightarrow$ $$\lim_{x\to a+0} \frac{f(x)}{g(x)} = \lim_{x\to a+0} \frac{f'(x)}{g'(x)}$$

For $x < a$, we can similarly prove that

$$\lim_{x\to a-0} \frac{f(x)}{g(x)} = \lim_{x\to a-0} \frac{f'(x)}{g'(x)}$$

Hence

$$\lim_{x\to a} \frac{f(x)}{g(x)} = \lim_{x\to a} \frac{f'(x)}{g'(x)}$$

It may be noted in the above theorem that if $g'(x) = 0$ at any point of the interval $]a-\delta, a+\delta[$, then $f'(x)$ also vanishes there at so that $f'(x)/g'(x)$ then takes up the indeterminate form 0/0. In such situations we have to proceed to the next two theorems which may be considered as generalisations of these theorems and may be proved by repeated applications of the above theorems.

Theorem 5. *Generalised L'Hospital's Rule for 0/0 form. If f, g be two functions such that*

(i) $f^n(x)$, $g^n(x)$ exist, and $g'(x) \neq 0$ $(r = 0, 1, 2, ..., n)$ for any x in $]a-\delta, a+\delta[$ except possibly at $x = a$,

(ii) when $x \to a$, $\begin{cases} \lim f(x) = \lim f'(x) = ... = \lim f^{n-1}(x) = 0 \\ \lim g(x) = \lim g'(x) = ... = \lim g^{n-1}(x) = 0 \end{cases}$

and (iii) $\lim_{x\to a} \dfrac{f^n(x)}{g^n(x)}$ exists, then

$$\lim_{x\to a} \frac{f(x)}{g(x)} = \lim_{x\to a} \frac{f^n(x)}{g^n(x)}$$

Theorem 6. *If f, g be two functions such that*

(i) when $x \to a$, $\begin{cases} \lim f(x) = \lim f'(x) = ... = \lim f^{n-1}(x) = 0 \\ \lim g(x) = \lim g'(x) = ... = \lim g^{n-1}(x) = 0 \end{cases}$

and (*ii*) $f^n(a)$, $g^n(a)$ *exist, and* $g^n(a) \neq 0$ *then*

$$\lim_{x \to a} \frac{f(x)}{g(x)} = \frac{f^n(a)}{g^n(a)}$$

[*Hint*: Using Theorem (3) it can be easily seen that

$$\lim_{x \to a} \frac{f^{n-1}(x)}{g^{n-1}(x)} = \frac{f^n(a)}{g^n(a)}$$

Also repeated applications of theorem (4) gives

$$\lim_{x \to a} \frac{f(x)}{g(x)} = \lim_{x \to a} \frac{f^{n-1}(x)}{g^{n-1}(x)}]$$

Remark: The rule in the form of theorem 4 or 5 is more useful than the other forms because we may cancel out common factors or affect any other simplification in the quotient $f'(x)/g'(x)$ before proceeding to the limits. Also repeated application of the rule is possible in this form.

L' Hospital's rule holds even in the case when each $f(x)$ and $g(x)$ tends to ∞ when $x \to a$ or when $x \to \infty$. Theorems 7 and 8 will show how these cases can be handled in an exactly similar fashion.

Example 5. If f' exists in the *nbd* of $x = a$ and $f''(a)$ exists. Show that

$$\lim_{h \to 0} \frac{f(a+2h) - 2f(a+h) + f(a)}{h^2}$$

exists and is equal to $f''(a)$. Give an example to show that the limit may exist even though $f'(a)$ does not exist.

■ $$\lim_{h \to 0} \frac{f'(a+2h) - f'(a)}{2h} = \lim_{h \to 0} \frac{f'(a+h) - f'(a)}{h} = f''(a)$$

Hence

$$\lim_{h \to 0} \frac{2f'(a+2h) - 2f'(a+h)}{2h} = \lim_{h \to 0} \left(2\left(\frac{f'(a+2h) - f'(a)}{2h} \right) - \left(\frac{f'(a+h) - f'(a)}{h} \right) \right)$$

$$= f''(a)$$

Then by theorem 5,

$$\lim_{h \to 0} \left(\frac{f(a+2h) - 2f(a+h) + f(a)}{h^2} \right)$$

exists and is equal to $f''(a)$.

Theorem 7. *L' Hospital's rule for infinite limits. If f, g be two functions such that*

(i) $\lim\limits_{x\to\infty} f(x) = \lim\limits_{x\to\infty} g(x) = 0$.

(ii) $f'(x)$, $g'(x)$ *exist, and* $g'(x) \neq 0$, $\forall\, x > 0$ *except possibly at* ∞, *and*

(iii) $\lim\limits_{x\to\infty} \dfrac{f'(x)}{g'(x)}$ *exists, then*

$$\lim_{x\to\infty} \frac{f(x)}{g(x)} = \lim_{x\to\infty} \frac{f'(x)}{g'(x)}$$

Let us put $z = 1/x$ so that $z \to 0+0$ as $x \to \infty$, and define two functions F and G where $F(z) = f(1/z)$ and $G(z) = g(1/z)$.

We see that the functions F and G satisfy the conditions of Theorem 4, viz.,

(i) $\lim\limits_{z\to 0+0} F(z) = \lim\limits_{z\to 0+0} G(z) = 0$,

(ii) $F'(z)$, $G'(z)$ exist and $G'(z) \neq 0$, $\forall\, z$ in $]-\delta, \delta[$ except possibly at $z = 0$, and

(iii) $\lim\limits_{z\to 0+0} \dfrac{F'(z)}{G'(z)}$ exists.

Consequently,

$$\lim_{z\to 0+0} \frac{F(z)}{G(z)} = \lim_{z\to 0+0} \frac{F'(z)}{G'(z)}$$

$$\Rightarrow \qquad \lim_{x\to\infty} \frac{f(x)}{g(x)} = \lim_{x\to\infty} \frac{f'(x)}{g'(x)}$$

2.3 Indeterminate Form, ∞/∞

If $f(x)$ and $g(x)$ both tend to ∞ as $x \to a$ then $\lim\limits_{x\to a} \dfrac{f(x)}{g(x)}$ takes ∞/∞ form. We shall prove L' Hospital's rule for this indeterminate form when x tends to a finite limit a. The rule, however, holds good even for infinite limits, $x \to \infty$ and may be deduced from the following theorem by the procedure indicated in Theorem 7.

Theorem 8[*]. *L' Hospital's rule for ∞/∞ form. If f, g be two functions such that*

(i) $\lim\limits_{x\to a} f(x) = \lim\limits_{x\to a} g(x) = \infty$,

(ii) $f'(x), g'(x)$ *exist and* $g'(x) \neq 0$, $\forall\, x$ *in* $]a-\delta, a+\delta[$, $\delta > 0$ *except possibly at a, and*

(iii) $\lim\limits_{x\to a} \dfrac{f'(x)}{g'(x)} = l$,

then

$$\lim_{x \to a} \frac{f(x)}{g(x)} = l$$

Consider the function

$$\phi(x) = f(x) - l g(x) \qquad \text{...(1)}$$

[Division by $g(x)$ would indicate that $f(x)/g(x) \to l$, if $\phi(x)/g(x) \to 0$. With this in mind we proceed to prove that the latter limit is actually equal to zero.]

Clearly, in view of condition (*ii*), $\phi(x)$ is derivable and continuous in $]a - \delta, a + \delta[$, except possibly at $x = a$. Also $g'(x) \neq 0$ there at.

$$\therefore \qquad \frac{\phi'(x)}{g'(x)} = \frac{f'(x)}{g'(x)} - l$$

But condition (*iii*) implies that $\exists$ a positive $\delta_1 < \delta$ and $\varepsilon > 0$ such that

$$\left| \frac{f'(x)}{g'(x)} - l \right| < \frac{\varepsilon}{2}, \text{ when } |x - a| < \delta_1$$

$$\therefore \qquad \left| \frac{\phi'(x)}{g'(x)} \right| < \frac{\varepsilon}{2}, \text{ for } |x - a| > \delta_1 \qquad \text{...(2)}$$

Let $x \neq a$ be a point of $]a - \delta_1, a + \delta_1[$.

For $x > a$, $\phi(x)$ and $g(x)$ satisfy all the conditions of Cauchy's Mean Value Theorem in $[x, a + \delta_1]$, so that

$$\frac{\phi(x) - \phi(a + \delta_1)}{g(x) - g(a + \delta_1)} = \frac{\phi'(\xi)}{g'(\xi)}, \text{ when } x < \xi < a + \delta_1 \qquad \text{...(3)}$$

Now in view of (2),

$$\left| \frac{\phi'(\xi)}{g'(\xi)} \right| < \frac{\varepsilon}{2}$$

Also since $g(x) \to \infty$ as $x \to a$, we have

(*i*) $\exists$ a positive $\delta_2 < \delta_1$ such that

$$\text{and} \qquad \left.\begin{array}{l} g(x) > 0 \\ g(x) > (2/\varepsilon)\,|\phi(a + \delta_1)| \end{array}\right\} \text{ when } (x - a) < \delta_2$$

(*ii*) $\exists$ a positive $\delta_3 < \delta_1$ such that

$$g(x) > g(a + \delta_1), \text{ when } (x - a) < \delta_3$$

Thus for $\delta_4 = \min(\delta_2, \delta_3)$, we have

$$\text{and} \qquad \left.\begin{array}{l} g(x) > 0 \\ g(x) - g(a + \delta_1) < g(x) \end{array}\right\} \text{ when } (x - a) < \delta_4$$

CHAPTER 7

Accordingly (3) gives

$$\frac{|\phi(x) - \phi(a + \delta_1)|}{g(x)} < \frac{\varepsilon}{2} \qquad \ldots(4)$$

Now,

$$\left|\frac{\phi(x)}{g(x)}\right| \le \left|\frac{\phi(x) - \phi(a + \delta_1)}{g(x)}\right| + \left|\frac{\phi(a + \delta_1)}{g(x)}\right|$$

$$< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon, \text{ for } (x - a) < \delta_4$$

Thus, $\lim\limits_{x\to a+0} \dfrac{\phi(x)}{g(x)} = 0$

For $x < a$, we can similarly prove that

$$\lim_{x\to a-0} \frac{\phi(x)}{g(x)} = 0$$

$\therefore$

$$\lim_{x\to a} \frac{\phi(x)}{g(x)} = 0$$

But since $\phi(x) = f(x) - lg(x)$, and $g(x) \neq 0$ for any $x \in]a - \delta_4, a + \delta_4[$ by (α), therefore dividing by $g(x)$ and taking limits,

$$\lim_{x\to a}\left[\frac{f(x)}{g(x)} - l\right] = \lim_{x\to a} \frac{\phi(x)}{g(x)} = 0$$

so that $\lim\limits_{x\to a} \dfrac{f(x)}{g(x)} = l$

Particular case. $l = 0$, *i.e.*, $\lim\limits_{x\to a} \dfrac{f'(x)}{g'(x)} = 0$.

The auxiliary function $\phi(x)$ reduces to $f(x)$, so that we have to proceed without introducing any such function. And we shall have the simplification, $\lim\limits_{x\to a} \dfrac{f'(x)}{g'(x)} = 0$, in place of $\lim\limits_{x\to a} \dfrac{\phi'(x)}{g'(x)} = 0$.

So proceeding in a similar way, we may prove the theorem:

If f, g be two functions such that

(*i*) $\lim\limits_{x\to a} f(x) = \lim\limits_{x\to a} g(x) = \infty$.

(*ii*) $f'(x), g'(x)$ *exist and* $g'(x) \neq 0$, *in a deleted neighbourhood of a, and*

(*iii*) $\lim\limits_{x\to a} \dfrac{f'(x)}{g'(x)} = 0$

then

$$\lim_{x\to a}\frac{f(x)}{g(x)}=0$$

Remarks:

1. $\lim\limits_{x\to a}\dfrac{f(x)}{g(x)}$ when $f(x)$ and $g(x)$ both tend to infinity, can be dealt with in the same way as 0/0 form.
2. The above theorem holds good in the case of infinite limits $x\to\infty$ as well.
3. Sometimes repeated applications of L'Hospital's rule may be necessary to evaluate a limit. We must then ensure at each step that the expression to which the rule is applied, is actually an indeterminate form.
4. The forms 0/0 and ∞/∞ can be interchanged and so care should be taken to select the form which would enable us to evaluate the limit most quickly.

2.4 (*a*) Form $0\times\infty$

When $f(x)\to 0$ and $g(x)\to\infty$ as $x\to a$, $f(x)\cdot g(x)$ takes $0\times\infty$ form.

However $f(x)\cdot g(x)$ may be expressed as

$$\frac{f(x)}{1/g(x)}\text{ or }\frac{g(x)}{1/f(x)}$$

which has respectively 0/0 and ∞/∞ forms.

(*b*) Form $\infty-\infty$

This can be reduced to the form 0/0 or ∞/∞. For

$$f(x)-g(x)=\frac{\dfrac{1}{g(x)}-\dfrac{1}{f(x)}}{\dfrac{1}{f(x)\,g(x)}}\qquad\left(\frac{0}{0}\text{ form}\right)$$

(*c*) Form $0^{\circ}, 1^{\infty}, \infty^{0}$

These forms can be made to depend upon one of the previous forms by putting $k=\{f(x)\}^{g(x)}$, so that

$$\log k = g(x)\cdot\log f(x)$$

$\therefore \lim\log k=\lim\,\{g(x)\log f(x)\}$

Also $\lim k=\lim e^{\log k}=e^{\lim\log k}$

Thus the limit may be evaluated by one of the previous methods.

2.5 Let us now evaluate some limits which take up these forms, we shall not hesitate to make use of certain known limits, such as $\lim\limits_{x\to 0}\dfrac{\sin x}{x}$, $\lim\limits_{x\to 0}\dfrac{\tan x}{x}$, $\lim\limits_{x\to\infty}\left(1+\dfrac{1}{x}\right)^{x}$ etc. or expansions of functions such as $\log(1+x)$, $\sin x$, etc. either in the beginning or at some intermediate stage because it simplifies and shortens the process of evaluation of a limit to a considerable extent.

Example 6. Evaluate $\lim\limits_{x\to 0} \dfrac{x - \tan x}{x^3}$.

■ Let $\dfrac{x - \tan x}{x^3} = \dfrac{f(x)}{g(x)}$, where $f(x) = x - \tan x$ and $g(x) = x^3$.

Now

$$\lim_{x\to 0} f(x) = \lim_{x\to 0} (x - \tan x) = 0$$

$$\lim_{x\to 0} g(x) = \lim_{x\to 0} x^3 = 0$$

Hence, $\lim\limits_{x\to 0} \dfrac{f(x)}{g(x)}$ is of $\dfrac{0}{0}$ form, so that L'Hospital's rule is applicable, *i.e.*,

$$\lim_{x\to 0} \frac{f(x)}{g(x)} = \lim_{x\to 0} \frac{f'(x)}{g'(x)}$$

$$\therefore \qquad \lim_{x\to 0} \frac{x - \tan x}{x^3} = \lim_{x\to 0} \frac{1 - \sec^2 x}{3x^2}$$

$$= -\frac{1}{3} \lim_{x\to 0} \left(\frac{\tan x}{x}\right)^2 = -\frac{1}{3}$$

Example 7. Evaluate $\lim\limits_{x\to 0} \dfrac{xe^x - \log(1+x)}{x^2}$.

■ It is a 0/0 form and therefore

$$\lim_{x\to 0} \frac{xe^x - \log(1+x)}{x^2} = \lim_{x\to 0} \frac{xe^x + e^x - \dfrac{1}{1+x}}{2x} \qquad \left(\frac{0}{0} \text{ form}\right)$$

$$= \lim_{x\to 0} \frac{xe^x + 2e^x + \dfrac{1}{(1+x)^2}}{2} = \frac{3}{2}$$

Example 8. Find $\lim\limits_{x\to 0} \dfrac{(1+x)^{1/x} - e}{x}$.

■ Since $\lim\limits_{x\to 0} (1+x)^{1/x} = e$, therefore it is a 0/0 form

$$\therefore \qquad \lim_{x\to 0} \frac{(1+x)^{1/x} - e}{x} = \lim_{x\to 0} \frac{\dfrac{d}{dx}(1+x)^{1/x}}{1}$$

$$= \lim_{x\to 0} \frac{(1+x)^{1/x}\{x-(1+x)\log(1+x)\}}{x^2(1+x)}$$

$$= e \cdot \lim_{x\to 0} \frac{x-(1+x)\log(1+x)}{x^2(1+x)} \qquad \left(\frac{0}{0} \text{ form}\right)$$

$$= e \cdot \lim_{x\to 0} \frac{-\log(1+x)}{2x+3x^2} \qquad \left(\frac{0}{0} \text{ form}\right)$$

$$= e \cdot \lim_{x\to 0} \frac{-1}{(2+6x)(1+x)} = -\frac{e}{2}$$

Aliter. Let us first find an algebraic expansion for $(1+x)^{1/x}$.

Let $y = (1+x)^{1/x}$

$$\log y = \frac{1}{x}\log(1+x)$$

$$= \frac{1}{x}\left(x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \ldots\right), \text{ if } |x| < 1$$

$$= 1 - \frac{x}{2} + \frac{x^2}{3} - \ldots$$

$$\therefore \qquad y = \exp\left(1 - \frac{x}{2} + \frac{x^2}{3} - \ldots\right)$$

$$= e \cdot \exp\left(-\frac{x}{2} + \frac{x^2}{3} - \ldots\right)$$

$$= e\left[1 + \left(-\frac{x}{2} + \frac{x^2}{3} - \ldots\right) + \frac{1}{2!}\left(-\frac{x}{2} + \frac{x^2}{3} - \ldots\right)^2 + \ldots\right]$$

$$= e\left[1 - \frac{x}{2} + \frac{11}{24}x^2 - \ldots\right]$$

$$\therefore \qquad \lim \frac{(1+x)^{1/x} - e}{x} = \lim_{x\to 0} \frac{e\left(1 - \frac{x}{2} + \frac{11}{24}x^2 - \ldots\right) - e}{x} = -\frac{e}{2}$$

Example 9. Find $\lim\limits_{x\to 1-0} \dfrac{\log(1-x)}{\cot(\pi x)}$.

■ It is a ∞/∞ form and therefore

$$\lim_{x\to 1-0} \frac{\log(1-x)}{\cot(\pi x)} = \lim_{x\to 1-0} \frac{\frac{-1}{1-x}}{-\pi \operatorname{cosec}^2(\pi x)}$$

$$= \lim_{x\to 1-0} \frac{\sin^2(\pi x)}{\pi(1-x)} \qquad \left(\frac{0}{0} \text{ form}\right)$$

$$= \lim_{x\to 1-0} \frac{\pi \sin(2\pi x)}{-\pi} = 0$$

Example 10. Evaluate $\lim_{x\to 0}\left(\frac{1}{x^2} - \frac{1}{\sin^2 x}\right)$.

■ It is a $(\infty - \infty)$ form, we therefore write as

$$\lim_{x\to 0}\left(\frac{1}{x^2} - \frac{1}{\sin^2 x}\right) = \lim_{x\to 0} \frac{\sin^2 x - x^2}{x^2 \sin^2 x} \qquad \left(\frac{0}{0} \text{ form}\right)$$

$$= \lim_{x\to 0} \frac{\sin 2x - 2x}{2x \sin^2 x + x^2 \sin 2x} \qquad \left(\frac{0}{0} \text{ form}\right)$$

$$= \lim_{x\to 0} \frac{\cos 2x - 1}{\left(\sin^2 x + 2x \sin 2x + x^2 \cos 2x\right)} \qquad \left(\frac{0}{0} \text{ form}\right)$$

$$= \lim_{x\to 0} \frac{-2 \sin 2x}{3 \sin 2x + 6x \cos 2x - 2x^2 \sin 2x} \qquad \left(\frac{0}{0} \text{ form}\right)$$

$$= \lim_{x\to 0} \frac{-\cos 2x}{3 \cos 2x - 4x \sin 4x - x^2 \cos 2x} = -\frac{1}{3}$$

Aliter.

$$\lim_{x\to 0}\left(\frac{1}{x^2} - \frac{1}{\sin^2 x}\right) = \lim_{x\to 0} \frac{\sin^2 x - x^2}{x^2 \sin^2 x}$$

$$= \lim_{x\to 0} \frac{\sin^2 x - x^2}{x^4}\left(\frac{x}{\sin x}\right)^2$$

$$\left[\text{Using } \lim_{x\to 0}\left(\frac{x}{\sin x}\right) = 1\right]$$

$$= \lim_{x\to 0} \frac{\sin^2 x - x^2}{x^4} \qquad \left(\frac{0}{0} \text{ form}\right)$$

$$= \lim_{x\to 0} \frac{\sin 2x - 2x}{4x^3} \qquad \left(\frac{0}{0} \text{ form}\right)$$

$$= \lim_{x\to 0} \frac{\cos 2x - 1}{6x^2}$$

$$= \lim_{x\to 0} \left[-\frac{1}{3}\left(\frac{\sin x}{x}\right)^2\right] = -\frac{1}{3}$$

Note: It should be noted that when a given function can be put as a product of two or more factors, the limit of each of which can be easily found, then limit of the entire function can be determined by evaluating the limit of each factor separately, provided that the product of these limits is not an indeterminate form.

Example 11. Evaluate $\lim_{x\to 0+0} (\sin x \log x)$.

■ It is a $(0 \times \infty)$ form. Let us write

$$\lim_{x\to 0+0} (\sin x \log x) = \lim_{x\to 0+0} \frac{\log x}{\operatorname{cosec} x} \qquad \left(\frac{\infty}{\infty} \text{ form}\right)$$

$$= -\lim_{x\to 0+0} \frac{1/x}{\operatorname{cosec} x \cot x} = -\lim_{x\to 0+0} \left(\frac{\sin x}{x}\right) \tan x = 0$$

Example 12. Evaluate $\lim_{x\to 0} \left(\frac{\tan x}{x}\right)^{1/x^2}$

■ It is a (1^∞) form. Therefore, let $K = \left(\frac{\tan x}{x}\right)^{1/x^2}$, so that

$$\log K = \frac{1}{x^2} \log\left(\frac{\tan x}{x}\right)$$

$$\therefore \qquad \lim_{x\to 0} \log K = \lim_{x\to 0} \frac{\log\left(\frac{\tan x}{x}\right)}{x^2} = \lim_{x\to 0} \frac{\frac{\sec^2 x}{\tan x} - \frac{1}{x}}{2x} \qquad \left(\frac{0}{0} \text{ form}\right)$$

$$= \lim_{x\to 0} \frac{x \sec^2 x - \tan x}{2x^2 \tan x} = \lim_{x\to 0} \frac{\sec^2 x \tan x}{2\tan x + x\sec^2 x} \qquad \left(\frac{0}{0} \text{ form}\right)$$

$$= \lim_{x\to 0} \frac{\tan x}{\sin 2x + x} = \lim_{x\to 0} \frac{\sec^2 x}{2\cos 2x + 1} = \frac{1}{3} \qquad \left(\frac{0}{0} \text{ form}\right)$$

i.e.,
$$\lim_{x\to 0} \log K = \frac{1}{3} \Rightarrow \lim_{x\to 0} K = e^{1/3}$$

$$\therefore \qquad \lim_{x\to 0} K = \lim_{x\to 0} \left(\frac{\tan x}{x}\right)^{1/x^2} = e^{1/3}$$

Example 13. Evaluate $\lim\limits_{x\to 1-0} (1 - x^2)^{1/(\log(1-x))}$.

■ It is a (0^0) form. Therefore, let $K = (1 - x^2)^{1/(\log(1-x))}$, so that

$$\log K = \frac{\log\left(1 - x^2\right)}{\log(1 - x)}$$

$$\therefore \qquad \lim_{x\to 1-0} \log K = \lim_{x\to 1-0} \frac{\log\left(1 - x^2\right)}{\log\left(1 - x\right)} = \lim_{x\to 1-0} \frac{2x(1 - x)}{1 - x^2} = \lim_{x\to 1-0} \frac{2x}{1 + x} = 1 \qquad \left(\frac{\infty}{\infty} \text{ form}\right)$$

$$\Rightarrow \qquad \lim_{x\to 1-0} \log K = 1 \Rightarrow \lim_{x\to 1-0} K = e.$$

EXERCISE

Evaluate the following limits:

1. $\lim\limits_{x\to 0} \dfrac{x - \sin x}{x^3}$

2. $\lim\limits_{x\to 0} \dfrac{x - \log(1 + x)}{1 - \cos x}$

3. $\lim\limits_{x\to 0} \dfrac{\log\left(1 + x^3\right)}{\sin^3 x}$

4. $\lim\limits_{x\to \infty} \dfrac{x^3}{e^x}$

5. $\lim\limits_{x\to \pi/2+0} \dfrac{\log\left(x - \dfrac{\pi}{2}\right)}{\tan x}$

6. $\lim\limits_{x\to \infty} x \tan \dfrac{1}{x}$

7. $\lim\limits_{x\to 0} \left(\cot^2 x - \dfrac{1}{x^2}\right)$

8. $\lim\limits_{x\to \pi/2} (\sec x - \tan x)$

9. $\lim\limits_{x\to 0+0} (\cot x)^{\sin x}$

10. $\lim\limits_{x\to \pi/2} (\sin x)^{\tan x}$

11. $\lim\limits_{x\to 0} \dfrac{(1 + x)^{1/x} - e + \dfrac{1}{2}ex}{x^2}$

    [*Hint*: Use algebraic expansion of $(1 + x)^{1/x}$.]

12. $\lim\limits_{x\to 0} \left(\dfrac{\sin x}{x}\right)^{1/x}$

13. $\lim_{x\to 0} \dfrac{e^x - 2\cos x + e^{-x}}{x \sin x}$

14. $\lim_{x\to 0} \left(\dfrac{1}{x} - \dfrac{1}{\sin x}\right)$

15. $\lim_{x\to 0} (1-x)^{1/x}$

16. $\lim_{x\to \infty} (1 + a/x)^x$, $a \neq 0$

17. $\lim_{x\to a} \left(2 - \dfrac{x}{a}\right)^{\tan(\pi x/2a)}$

18. $\lim_{x\to 1} x^{1/(x-1)}$

19. Find the values of a and b in order that $\lim_{x\to 0} \dfrac{x(1 + a\cos x) - b\sin x}{x^3}$ may be equal to 1.

20. Find the values of a and the limit in order that $\lim_{x\to 0} \dfrac{\sin 2x + a\sin x}{x^3}$ be finite.

21. If $f''(x)$ exists and is continuous in a neighbourhood of $x = a$, then show that

$$\lim_{h\to 0} \frac{f(a+h) - 2f(a) + f(a-h)}{h^2} = f''(a).$$

ANSWERS

1. $\frac{1}{6}$ 2. 1 3. 1 4. 0 5. 0 6. 1

7. $-\frac{2}{3}$ 8. 0 9. 1 10. 1 11. $\frac{1}{2}\frac{1}{4}e$ 12. 1

13. 2 14. 0 15. e^{-1} 16. e^a 17. $e^{2/\pi}$ 18. e

19. $a = -5/2$, $b = -3/2$ 20. $a = -2$, limit $= -1$.

8 Functions

In this chapter, we propose to introduce the so-called Elementary functions.

$$e^x,\ \log x,\ a^x,\ \sin x,\ \cos x$$

The reader is already familiar with these functions but this acquaintance is based on a treatment which was essentially based on intuitive and less rigorous geometrical considerations. Even the question of existence was ignored.

We shall base the study of these functions on the set of real numbers as a complete ordered field, the notion of limit and the convergence of series. Starting from the definitions of these functions, their basic properties will be studied. No attempt, however, will be made to make the discussion exhaustive.

We shall consciously accept an abuse of language in as much as notation for a function will not be distinguished from that for the functional value. Thus instead of denoting the function by cos we shall denote the same as $\cos x$ and so on.

Functions of bounded variation and the vector-valued functions have been considered towards the end.

As the definitions of function will be based on *Power series*, we start our discussion with a brief (very brief!) study of power series. A detailed discussion will be found in Chapter 14.

1. POWER SERIES

The series of the form

$$a_0 + a_1 x + a_2 x^2 + \dots + a_n x^n + \dots = \sum_{n=0}^{\infty} a_n x^n \text{ or simply } \Sigma a_n x^n$$

are called *power series* (in x) and the numbers a_n are their *coefficients*.

For such series we are not concerned simply with the alternatives 'convergent' and 'divergent', but the more precise question: For what values of x is the series convergent and for what values is divergent?

1.1 Some simple examples have already come before us:

1. The geometric series Σx^n is convergent for $|x| < 1$, divergent for $|x| \geq 1$.

   For $|x| < 1$, indeed, we have absolute convergence.

2. $\Sigma \frac{x^n}{n!}$ is (absolutely) convergent for every real x; likewise the series $\Sigma(-1)^n \frac{x^{2n}}{2n!}$ and $\Sigma(-1)^n \frac{x^{2n+1}}{(2n+1)!}$.

3. $\Sigma \frac{x^n}{n}$.

   For $x = 1$, diverges (reduces to divergent harmonic series)
   $x = -1$, converges
   $|x| < 1$ absolutely convergent ($\leq |x|^n$, comparison test)
   $|x| > 1$ diverges

4. $\Sigma n^n x^n$ converges for $x = 0$, but diverges for $x \neq 0$.

1.2 For $x = 0$, obviously every power series $\Sigma a_n x^n$ is convergent, whatever be the value of the coefficient a_n. The general case is evidently that in which the power series converges for some values of x and diverges for others; while in special instances, the two extreme cases may occur, in which the series converges for every x or for none $x \neq 0$.

In the first of these special cases, when the series converges for every x, we say that the power series is *everywhere convergent*, while in the second (leaving out of account the self-evident point of convergence $x = 0$) when the series converges for no value of x we say that it is *nowhere convergent*. The totality of points x for which the series converges is called its *region of convergence*.

For a power series $\Sigma a_n x^n$, which does not merely converge everywhere or nowhere, a definite positive number R exists such that the series converges for every $|x| < R$ (indeed absolutely), but diverges for every $|x| > R$. The number R is called the *Radius of convergence*, and the interval $]-R, R[$, the *Interval of convergence* of the given series. The behaviour of the series is much more varied at $x = \pm R$ and is beyond the scope of the present discussion.

1.3 For the power series $\Sigma a_n x^n$, put

$$\lim_{n \to \infty} \sqrt[n]{|a_n|} = \alpha \text{ and } R = \frac{1}{\alpha}$$

where $R = +\infty$, for $\alpha = 0$ and $R = 0$, for $\alpha = \infty$.

By Cauchy's Root test, it follows that if

(*i*) $R = 0$, the series is nowhere convergent;

(*ii*) $R = \infty$, the series is everywhere convergent;

(*iii*) $0 < R < \infty$, the series converges absolutely for $|x| < R$, and diverges for $|x| > R$, *i.e.*, R is the radius of convergence.

Ex. 1. $1 + x + x^2 + \ldots$ converges for $|x| < 1$ and equals $(1 - x)^{-1}$.

CHAPTER 8

2. EXPONENTIAL FUNCTIONS

The power series

$$1+\frac{x}{1!}+\frac{x^2}{2!}+\frac{x^3}{3!}+...+\frac{x^n}{n!}+... \qquad ...(1)$$

is everywhere convergent for real x. We proceed now to examine in detail the function represented by this series.

2.1 Definition

The function represented by the power series (1) is called the *Exponential function*, denoted, provisionally, by $E(x)$. Thus

$$E(x)=1+\frac{x}{1!}+\frac{x^2}{2!}+...+\frac{x^n}{n!}+... \qquad ...(2)$$

$$\therefore \qquad E(0)=1$$

and

$$E(1)=1+\frac{1}{1!}+\frac{1}{2!}+...+\frac{1}{n!}+... \qquad ...(3)$$

The series on the right hand side of (3) converges to a number which lies between 2 and 3. This number is denoted by e, the *Exponential base* and is the same number as represented by

$$\lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n$$

Thus $E(1)=e$

2.2 The Additional Formula

The function $E(x)$, defined by (2), is continuous and differentiable any number of times, for every x.

By differentiation, we get

$$E'(x)=E(x)$$

$$E''(x)=E(x)$$

$$\cdots\cdots\cdots\cdots$$

$$E^{(n)}(x) = E(x)$$

Further we state (justification may be seen expanding by Taylor's Theorem)* that

$$E(x_1 + x_2) = E(x_1).E(x_2)$$

---

* $E(x)=E(x_1)+\dfrac{E(x_1)}{1!}(x-x_1)+...$, for all values of x and x_1.

Replacing x by x_1+x_2, we get

$$E(x_1+x_2)=E(x_1)\left\{1+\frac{x_2}{1!}+\frac{x_2^2}{2!}+...\right\}=E(x_1)\cdot E(x_2)$$

This formula is called the *Addition formula* for the exponential function. It gives further

$$\begin{aligned} E(x_1 + x_2 + x_3) &= E(x_1 + x_2) \cdot E(x_3) \\ &= E(x_1) \cdot E(x_2) \cdot E(x_3) \end{aligned}$$

and repetition of the process gives, for any positive integer q,

$$E(x_1 + x_2 + ... + x_q) = E(x_1) \cdot E(x_2) ... E(x_q) \quad ...(4)$$

If $x_1 = x_2 = x_3 = ... = x_q = x$, we get

$$E(qx) = \{E(x)\}^q \quad ...(5)$$

Hence, for $x = 1$

$$E(q) = \{E(1)\}^q = e^q, \text{ for any positive integer } q$$

But since $E(0) = 1$, therefore the above relation holds for $q = 0$ also.

Hence $E(q) = e^q$ holds for all integers ≥ 0.

Again replacing each x by p/q in (5), we get

$$E\left(q\frac{p}{q}\right) = \left\{E\left(\frac{p}{q}\right)\right\}^q \text{ for positive integers, } p, q$$

or

$$E(p/q) = \{E(p)\}^{1/q} = e^{p/q} \qquad \left[\because\ E(p) = e^p\right]$$

Hence $E(m) = e^m$, for all rational numbers $m \geq 0$.

For any positive *irrational* number ξ there always exists a sequence $\{x_n\}$ of positive rational terms, converging to ξ.

Now for each n

$$E(x_n) = e^x n.$$

When $n \to +\infty$, the left hand side tends to $E(\xi)$, and the right hand side to e^ξ, so that we get

$$E(\xi) = e^\xi$$

$$\therefore \qquad E(x) = e^x, \text{ for real } x \geq 0 \quad ...(6)$$

Again by Addition formula,

$$E(x) \cdot E(-x) = E(x - x) = E(0) = 1 \quad ...(7)$$

Thus we conclude that $E(x) \neq 0$, for any real x, and that for $x \geq 0$,

$$E(-x) = \frac{1}{E(x)} = \frac{1}{e^x} = e^{-x},$$

Consequently, $E(x) = e^x$ holds for all real x.

2.3 Monotonicity

By definition,

$$E(x) > 0, \ \forall\, x > 0$$

so that from (7) it follows that

$$E(-x) > 0, \ \forall\, x > 0$$

Hence, $E(x) > 0$, for all real x.

Again by definition, for real x,

$$E(x) \to +\infty, \text{ as } x \to +\infty$$

Hence, (7) shows that

$$E(x) \to 0 \text{ as } x \to -\infty$$

Also by definition,

$$0 < x_1 < x_2 \Rightarrow E(x_1) < E(x_2)$$

Also it follows from (7) that

$$E(-x_2) < E(-x_1), \text{ when } -x_2 < -x_1 < 0$$

Hence, the function E is strictly increasing from 0 to $+\infty$ on the whole real line.

Note: By definition $e^x > \dfrac{x^{n+1}}{(n+1)!}$, for $x > 0$, so that $x^n e^{-x} < \dfrac{(n+1)!}{x}$.

$\therefore \lim\limits_{x \to +\infty} x^n e^{-x} = 0$, for all n

This fact we express by saying that e^x *tends to* $+\infty$ *"faster" than any power* of x, as $x \to +\infty$.

3. LOGARITHMIC FUNCTIONS (*base e*)

Since the exponential function E is strictly increasing on the set **R** of real numbers (*i.e.*, E: $\mathbf{R} \to \mathbf{R}^+$ is one-one onto), it has inverse function L (or $\log_e$) which is also strictly increasing and whose domain of definition is $\mathbf{R}^+$ $(\equiv E(\mathbf{R}))$, the set of positive reals. Thus **L** is defined by

or

$$\left.\begin{array}{l} E\{L(y)\} = y, \ (y > 0) \\ L\{E(x)\} = x, \ (x \text{ real}) \end{array}\right\} \qquad \ldots(i)$$

or equivalently, for any real x,

or

$$\left.\begin{array}{r} E(x) = y \Rightarrow L(y) = x \\ e^x = y \Rightarrow \log_e y = x \end{array}\right\} \qquad \ldots(ii)$$

Thus the logarithmic function L (or $\log_e$) is defined for positive values only of the variable.

By definition,

$$\left.\begin{array}{lll} E(-x) = \dfrac{1}{y} & \Rightarrow & L\left(\dfrac{1}{y}\right) = -x = -L(y) \\ E(0) = 1 & \Rightarrow & L(1) = 0 = \log_e 1 \\ E(1) = e & \Rightarrow & L(e) = 1 = \log_e e \end{array}\right\} \qquad \ldots(iii)$$

Again

$\because$ $$E(x) \to +\infty \text{ as } x \to +\infty$$

and

$$E(x) \to 0 \text{ as } x \to -\infty$$

$\therefore$ $$L(x) \to +\infty \text{ as } x \to +\infty$$

$$L(x) \to -\infty \text{ as } x \to 0$$

Writing $u = E(x_1)$, $v = E(x_2)$ or $L(u) = x_1$, $L(v) = x_2$ in (4), we get

$$E(x_1 + x_2) = uv$$

$\Rightarrow$ $$L(uv) = x_1 + x_2 = L(u) + L(v)$$

which is a familiar property of the logarithmic function and which makes logarithms a useful tool for computation.

Since, the function E is differentiable, therefore, its inverse function **L** is also differentiable.

Hence differentiating (*i*), we get

$$L'\{E(x)\} \cdot E(x) = 1$$

Writing $E(x) = y$, we get

$$L'(y) = \frac{1}{y} \qquad ...(iv)$$

which implies that

$$L(y) = \int_1^y \frac{dx}{x} \qquad ...(8)$$

Quite often (8) is taken as the definition of the logarithmic function and thus the starting point of the theory of the logarithmic and the exponential functions.

Note: In theoretical investigations, it is always more convenient to use the so-called *natural logarithms*, that is to say, those with the base e. Hence in our further discussion, $\log x$ *shall always stand for* $L(x)$ or $\log_e x$.

3.1 Generalised Power Functions

The meaning of a^x is well understood when a is any positive real number and x is any rational number. We shall now give a meaning to a^x when x is any real number whatsoever. We define thus:

Definition. $a^x = E(x \log a)$, for all x and $a > 0$.

Evidently the range of a^x is the set $\mathbf{R}^+$ of positive reals, *i.e.*,

$$a^x > 0, \ \forall x \qquad ...(9)$$

$$\text{Now } a^x \cdot a^y = E(x \log a) \cdot E(y \log a)$$
$$= E\{(x+y) \log a\} = a^{x+y}$$

$\therefore$ $$a^x \cdot a^y = a^{x+y} \qquad ...(10)$$

Let us now verify that this definition of a^x is consistent with that already known to us for x, an integer or a rational number.

CHAPTER 8

(*i*) Let $x = n$, a positive integer.

$$\therefore \quad a^n = E(n \log a) = E[\log a + \log a + \ldots n \text{ times}]$$
$$= E(\log a) \cdot E(\log a) \ldots n \text{ times}$$
$$= a \cdot a \ldots n \text{ times}$$

(*ii*) Now let $x = -n$, n being a positive integer.

$$\therefore \quad a^{-n} = E(-n \log a)$$
$$= E[(-\log a) + (-\log a) + \ldots n \text{ times}]$$
$$= E\left[\log \frac{1}{a} + \log \frac{1}{a} + \ldots n \text{ times}\right]$$
$$= E\left(\log \frac{1}{a}\right) \cdot E\left(\log \frac{1}{a}\right) \ldots n \text{ times}$$
$$= \frac{1}{a} \cdot \frac{1}{a} \ldots n \text{ times}$$

Thus $E(x \log a)$ has the same meaning as a^x when x is an integer.

(*iii*) Let now $x = p/q$, where p, q are integers, and q is positive.

Now

$$E\left(\frac{p}{q} \log a\right) = a^{p/q}$$

$$\therefore \quad \left[E\left(\frac{p}{q} \log a\right)\right]^q = a^p = E(p \log a)$$

so that $E\left(\frac{p}{q} \log a\right)$ is qth root of $E(p \log a)$.

Thus $a^{p/q}$ is a qth root of a^p.

Hence the definition holds good when x is a rational number.

Thus the above definition of a^x agrees with what is already known to us about a^x.

3.2 Logarithmic Functions (*any base*)

Definition. $a^x = y \quad \Leftrightarrow \quad \log_a y = x$.

Since y is always positive, therefore the logarithmic function, $\log_a$, is defined for positive values only of the variable.

Evidently

$$a^{-x} = \frac{1}{y}$$

$$\therefore \quad \log_a \frac{1}{y} = -x = -\log_a y$$

Also, from definition,

$$\log_a 1 = 0, \log_a a = 1$$

It may be easily shown that

$$\log_a x + \log_a y = \log_a (xy)$$
$$\log_a x - \log_a y = \log_a (x/y)$$
$$\log_a x^y = y \log_a x$$
$$\log_b x \cdot \log_a b = \log_a x$$
$$\log_b a \cdot \log_a b = 1$$

Ex. Show that $\lim\limits_{x\to\infty} \dfrac{\log x}{x^a} = 0,\ a > 0.$

[*Hint*: Use $\lim\limits_{n\to\infty} x^n/e^x = 0.$]

4. TRIGONOMETRIC FUNCTIONS

We are now in a position to introduce rigorously the circular functions, employing purely the arithmetical methods. For this purpose, we consider the power series, everywhere convergent (absolutely and uniformly) and the functions represented by them.

Definition.

$$C(x) = 1 - \frac{x^2}{2!} + \frac{x^2}{2!} - \ldots + (-1)^n \frac{x^{2n}}{(2n)!} + \ldots \ \forall\, x$$

$$S(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \ldots + (-1)^n \frac{x^{2n+1}}{(2n+1)!} + \ldots \ \forall\, x$$

Each of these series represents a function everywhere continuous and differentiable any number of times in succession. The properties of these functions will be established, taking as starting point their expansions in series form, and it will be seen finally that these coincide with the functions cos x and sin x with which we are familiar from elementary studies, *i.e.*, $C(x) \equiv \cos x$ and $S(x) \equiv \sin x$.

4.1 Properties of the Functions ($C(x)$, $S(x)$)

(*i*) The functions $C(x)$ and $S(x)$ are continuous and derivable for all x; in fact it may easily be seen that

$$C'(x) = -S(x) \text{ and } S'(x) = C(x)$$

Also

$$C^{iv}(x) = C(x) \text{ and } S^{iv}(x) = S(x)$$

(*ii*) From definitions,

$$S(0) = 0, C(0) = 1$$

$$S(-x) = -x - \frac{(-x)^3}{3!} + \frac{(-x)^5}{5!} - \ldots + (-1)^n \frac{(-x)^{2n+1}}{(2n+1)!} + \ldots$$

CHAPTER 8

$$= -\left[x - \frac{x^3}{3!} + \frac{x^5}{5!} - \dots + (-1)^n \frac{x^{2n+1}}{(2n+1)!} + \dots\right] = -S(x) \quad \forall\, x$$

Similarly, $C(-x) = C(x) \quad \forall\, x$

(iii) **The Addition Theorems.** These functions, like the exponential function, satisfy simple addition theorems, by means of which they can then be further examined.

First Method. By Taylor's expansion for any two variables, x_1 and x_2 (since the two series converge everywhere absolutely).

$$C(x_1 + x_2) = C(x_1) + \frac{C'(x_1)}{1!}x_2 + \frac{C''(x_1)}{2!}x_2^2 + \dots$$

$$= C(x_1) - \frac{S(x_1)}{1!}x_2 - \frac{C(x_1)}{2!}x_2^2 + \frac{S(x_1)}{3!}x_2^3 + \dots$$

As this series is absolutely convergent, we may rearrange it in any way we please.

$$\therefore\ C(x_1 + x_2) = C(x_1)\left\{1 - \frac{x_2^2}{2!} + \frac{x_2^4}{4!} - \dots\right\} - S(x_1)\left\{x_2 - \frac{x_2^3}{3!} + \frac{x_2^5}{5!} - \dots\right\}$$

$$= C(x_1) \cdot C(x_2) - S(x_1) \cdot S(x_2) \qquad \dots(11)$$

Similarly,

$$S(x_1 + x_2) = S(x_1) \cdot C(x_2) + C(x_1) \cdot S(x_2) \qquad \dots(12)$$

Second Method. For any fixed value of x_2, consider the functions

$$f(x_1) = S(x_1 + x_2) - S(x_1) \cdot C(x_2) - C(x_1) \cdot S(x_2)$$
$$g(x_1) = C(x_1 + x_2) - C(x_1) \cdot C(x_2) + S(x_1) \cdot S(x_2)$$

Differentiating with respect to x_1, we get

$$f'(x_1) = C(x_1 + x_2) - C(x_1) \cdot C(x_2) + S(x_1) \cdot S(x_2) = g(x_1)$$
$$g'(x_1) = -S(x_1 + x_2) + S(x_1) \cdot C(x_2) + C(x_1) \cdot S(x_2) = -f(x_1)$$

$$\therefore \qquad \frac{d}{dx_1}\left[f^2(x_1) + g^2(x_1)\right] = 2f(x_1)f'(x_1) + 2g(x_1)g'(x_1)$$

$$= 2f(x_1)\,g(x_1) - 2g(x_1)f(x_1) = 0, \quad \forall\, x_1$$

$\Rightarrow \quad f^2(x_1) + g^2(x_1)$ is a constant, $\forall\, x_1$

Hence for all x_1,

$$f^2(x_1) + g^2(x_1) = f^2(0) + g^2(0) = 0$$

$$\Rightarrow \qquad f(x_1) = 0,\ g(x_1) = 0$$

$$\therefore \qquad C(x_1 + x_2) = C(x_1) \cdot C(x_2) - S(x_1) \cdot S(x_2)$$

and

$$S(x_1 + x_2) = S(x_1) \cdot C(x_2) + C(x_1) \cdot S(x_2)$$

The form of these theorems coincides with that of the addition theorems for the functions cosine and sine, with which we are clearly acquainted from an elementary standpoint. With the help of these theorems, we shall now show that the functions C and S satisfy all the other so called purely trigonometrical formulae—in fact C and S are same as the functions cosine and sine. We note, in particular:

(*a*) Changing x_2 to $-x_2$,

$$C(x_1 - x_2) = C(x_1) \cdot C(x_2) + S(x_1) \cdot S(x_2)$$
$$S(x_1 - x_2) = S(x_1) \cdot C(x_2) - C(x_1) \cdot S(x_2)$$

(*b*) Writing $x_2 = -x_1$, we deduce that

$$C^2(x_1) + S^2(x_1) = 1 \text{ or } C^2(x) + S^2(x) = 1, \ \forall x$$
$$\Rightarrow |S(x)| \le 1, |C(x)| \le 1, \ \forall x$$

(*c*) Replacing x_1 and x_2 by x,

$$C(2x) = C^2(x) - S^2(x)$$
$$S(2x) = 2S(x) \cdot C(x)$$

4.2 The Number π — The Smallest Positive Root of the Equation $C(x) = 0$.

Theorem 1. *To prove that there exists a positive number π such that*

$$C(\pi/2) = 0 \text{ and } C(x) > 0, \text{ for } 0 \le x < \pi/2 \qquad \text{...(13)}$$

Consider the interval [0, 2].

We know $C(0) = 1 > 0$, we shall now show that $C(2) < 0$.

Now

$$C(2) = 1 - \frac{2^2}{2!} + \frac{2^4}{4!} - \frac{2^6}{6!} + \ldots$$
$$= 1 - \frac{2^2}{2!}\left(1 - \frac{2^2}{3.4}\right) - \frac{2^6}{6!}\left(1 - \frac{2^2}{7.8}\right) - \ldots$$

Since the brackets are all positive, we have

$$C(2) < 1 - \frac{2^2}{2!}\left(1 - \frac{2^2}{3.4}\right) = -\frac{1}{3}$$

so that $C(2)$ is negative.

Thus, the continuous function $C(x)$ is positive at 0 and negative at 2.

$\therefore$ $C(x)$ vanishes at least once between 0 and 2 (by the Intermediate-value theorem). Further, since $S(x)$ is positive in [0, 2], where

$$S(x) = x\left(1 - \frac{x^2}{2.3}\right) + \frac{x^5}{5!}\left(1 - \frac{x^2}{6.7}\right) + \ldots$$

therefore, the derivative $(-S(x))$ of $C(x)$ is always negative for all values of x between 0 and 2. Consequently $C(x)$ is a (strictly) monotonic decreasing function in [0, 2], and can therefore vanish at only one point in [0, 2].

Thus there exists one and only one root of the equation $C(x) = 0$ lying between 0 and 2. Denoting this root by $\pi/2$, we see that $\pi/2$ is the least positive root of the equation $C(x) = 0$.

Clearly $C(x) > 0$, when $0 \le x < \pi/2$.

Using the above results, we deduce that

(*a*) $S(x) > 0$, when $0 < x \le \pi/2$.

Since the derivative of $S(x)$ is non-negative in $[0, \pi/2]$, therefore $S(x)$ is a strictly monotonic increasing function. Also since $S(0) = 0$, therefore $S(x)$ is positive for $0 < x \le \pi/2$.

(*b*) As $C^2(\pi/2) + S^2(\pi/2) = 1$ and $C(\pi/2) = 0$,

$$\therefore \quad S^2(\pi/2) = 1 \Rightarrow S(\pi/2) = \pm 1$$

But, by Lagrange's Mean Value Theorem,

$$S(\pi/2) - S(0) = (\pi/2)C(\alpha) > 0, \text{ where } 0 < \alpha < \pi/2$$

$$\Rightarrow \quad S(\pi/2) = 1$$

(*c*) $C(\pi) = 2C^2(\pi/2) - 1 = -1$

$$S(\pi) = 2S(\pi/2)\,C(\pi/2) = 0$$

(*d*) $C(2\pi) = 1,\ S(2\pi) = 0$

(*e*) $C(\pi/2) = 2C^2(\pi/4) - 1$

$$\therefore \quad C(\pi/4) = 1/\sqrt{2}$$

Rejecting the negative sign, as $C(\pi/4)$ is positive.

Similarly, $S(\pi/4) = 1/\sqrt{2}$

(*f*) It finally follows from the addition theorems that for all x,

$$S\left(\frac{1}{2}\pi - x\right) = C(x),\quad C\left(\frac{1}{2}\pi - x\right) = S(x)$$

$$S\left(\frac{1}{2}\pi + x\right) = C(x),\quad C\left(\frac{1}{2}\pi + x\right) = -S(x)$$

$$S(\pi + x) = -S(x),\quad C(\pi + x) = -C(x)$$

$$S(\pi - x) = S(x),\quad C(\pi - x) = -C(x)$$

$$S(2\pi + x) = S(x),\quad C(2\pi + x) = C(x)$$

Thus, we see that the functions $C(x)$ and $S(x)$ exactly coincide with the functions $\cos x$ and $\sin x$ respectively, and so we shall henceforth use $\cos x$ and $\sin x$ in place of $C(x)$ and $S(x)$ respectively.

4.3 The Functions tan *x* and cot *x*

The function $\tan x$ and $\cot x$ are defined as usual by the ratios:

$$\tan x = \frac{\sin x}{\cos x},\ \cot x = \frac{\cos x}{\sin x}$$

and as functions they, therefore, represent nothing new. The expansions in power series for these functions are also not so simple. A few of the coefficients of the expansions could be easily obtained by division, but that gives us no insight into any relationships.

Clearly tan x is defined, continuous and derivable for all values of x except those for which the denominator, cos x, vanishes, which is the case for $x = \frac{1}{2}(2n+1)\pi$, n being any integer, positive, negative or zero.

From § 4.2 (*f*), we have

$$\tan(\pi + x) = \tan x,$$

so that, tan x is a periodic function with period π.

Also we may easily show that when $x \neq \frac{1}{2}(2n+1)\pi$,

$$\frac{d}{dx}\tan x = \frac{d}{dx}\left(\frac{\sin x}{\cos x}\right) = \frac{1}{\cos^2 x}$$

Theorem 2. *Show that*

$$\lim_{x \to \frac{1}{2}\pi - 0} \tan x = \infty, \quad \lim_{x \to \frac{1}{2}\pi + 0} \tan x = -\infty$$

Let k be any positive number.

As $\lim\limits_{x \to \pi/2} \sin x = 1$, $\exists\, \delta_1 > 0$ such that (taking $\varepsilon = \frac{1}{2}$),

$$\frac{1}{2} < \sin x, \quad \forall\, x \in \left]\frac{1}{2}\pi - \delta_1, \frac{1}{2}\pi + \delta_1\right[\qquad ...(i)$$

Again, since $\lim\limits_{x \to \pi/2} \cos x = 0$, therefore $\exists\, \delta_2 > 0$ such that

$$-\frac{1}{2k} < \cos x < \frac{1}{2k}, \quad \forall\, x \in \left]\frac{1}{2}\pi - \delta_2, \frac{1}{2}\pi + \delta_2\right[$$

As $\cos x$ is positive for $x \in [0, \pi/2[$, and negative for $x \in]\pi/2, \pi]$, we have

$$0 < \cos x < \frac{1}{2k}, \quad \forall\, x \in \left]\frac{1}{2}\pi - \delta_2, \frac{1}{2}\pi\right[\qquad ...(ii)$$

$$-\frac{1}{2k} < \cos x < 0, \quad \forall\, x \in \left]\frac{1}{2}\pi, \frac{1}{2}\pi + \delta_2\right[\qquad ...(iii)$$

Let $\delta = \min(\delta_1, \delta_2)$ therefore from (*i*) and (*ii*),

$$\tan x = \frac{\sin x}{\cos x} > k, \quad \forall\, x \in \left]\frac{1}{2}\pi - \delta, \frac{1}{2}\pi\right[$$

and from (*i*) and (*iii*),

$$\tan x = \frac{\sin x}{\cos x} < -k, \quad \forall\, x \in \left]\frac{1}{2}\pi, \frac{1}{2}\pi + \delta\right[$$

CHAPTER 8

4.4 Inverse Trigonometric Functions $\cos^{-1} y$, $\sin^{-1} y$, $\tan^{-1} y$

$\cos^{-1} y$. Since, as may be easily seen, $\cos x$ strictly decreases from $+1$ to -1 as x increases from 0 to π, the function cos is invertible and its inverse, denoted as $\cos^{-1}$, is a function with domain $[-1, 1]$ and range $[0, \pi]$. We write

$$y = \cos x \quad \Leftrightarrow \quad x = \cos^{-1} y.$$

Definition. Given y (where $-1 \le y \le 1$), $\cos^{-1} y$ is that x which lies between 0 and $\pi\left(0 \le x \le \pi\right)$ and $\cos x = y$.

$\cos^{-1} y$ is derivable in the open interval $]-1, 1[$ with $-1/\sqrt{1-y^2}$ as its derivative. In fact, we have

$$\frac{dx}{dy} \cdot \frac{dy}{dx} = 1, \text{ and } x = \cos^{-1} y,\ y = \cos x$$

$$\frac{d}{dy}\left(\cos^{-1} y\right) = \frac{1}{\dfrac{d}{dx}\cos x} = -\frac{1}{\sin x} = \frac{-1}{\sqrt{1-y^2}},\ y \ne \pm 1$$

$\sin^{-1} y$. Since $\sin x$ is a strictly increasing function in $\left[-\pi/2, \pi/2\right]$ with range $[-1, 1]$, therefore the function sin is invertible and its inverse function is denoted by $\sin^{-1}$, with domain $[-1, 1]$ and range $\left[-\pi/2, \pi/2\right]$.

Also

$$y = \sin x \Leftrightarrow x = \sin^{-1} y$$

Definition. Given y where $-1 \le y \le 1$, $\sin^{-1} y$ is that x which lies between $-\pi/2$ and $\pi/2$, $\left(-\pi/2 \le x \le \pi/2\right)$ and $\sin x = y$.

It may be shown as before that $\sin^{-1} y$ is derivable in the open interval $]-1, 1[$ and

$$\frac{d}{dy}\left(\sin^{-1} y\right) = \frac{1}{\sqrt{1-y^2}},\ y \ne \pm 1$$

$\tan^{-1} y$. Since $\tan x$ is strictly monotonic with domain $\left]-\pi/2, \pi/2\right[$ and range $\left]-\infty, \infty\right[$, the function is invertible, we have

$$y = \tan x \quad \Leftrightarrow \quad x = \tan^{-1} y$$

so that $\tan^{-1} y$ is a function with domain $\left]-\infty, \infty\right[$ and range $\left]-\pi/2, \pi/2\right[$.

Definition. For any number y, $\tan^{-1} y$ is that x which lies between $-\pi/2$ and $\pi/2\left(-\pi/2 < x < \pi/2\right)$ and $\tan x = y$.

It may be seen that

$$\frac{d}{dy}\tan^{-1} y = \frac{1}{1+y^2},\ \forall\ y$$

5. FUNCTIONAL EQUATIONS

In this section we shall discuss the solutions of certain functional equations which, together with continuity, suffice to characterize the so-called elementary functions.

First we note that the continuous function $f(x) = cx$, $x \in \mathbf{R}$, c being constant, satisfies the *functional equation*,

$$f(x+y) = f(x) + f(y), \ \forall\, x, y \in \mathbf{R} \qquad ...(1)$$

The interesting fact is that *every continuous function satisfying the functional equation* (1) *is of this form*.

In order to prove this, let $f : \mathbf{R} \to \mathbf{R}$ be continuous and satisfies the functional equation (1). Then we have,

$$f(0) = 0, \text{ and } f(-x) = -f(x), \text{ for all } x \in \mathbf{R}$$

Also, for each positive integer n, we have by induction on n,

$$f(nx) = nf(x), \text{ for all } x \in \mathbf{R}$$

Replacing x by x/n in above, we obtain

$$f(x/n) = (1/n)\, f(x), \text{ for all } x \in \mathbf{R}$$

Thus, for any pair of integers p, q (q being positive), we have

$$f\left(\frac{p}{q}x\right) = \frac{p}{q} f(x), \text{ for all } x \in \mathbf{R}$$

In other words, $f(rx) = rf(x)$, for any rational number r.

If we put $x = 1$ and $f(1) = c$ in above, then we obtain

$f(r) = cr$, for every rational number r.

Now, let ξ be any *irrational* number, then there always exists a sequence $\{r_n\}$ of rational numbers, converging to ξ.

$$\therefore \qquad f(\xi) = \lim_{x\to\infty} f(r_n) = \lim_{x\to\infty} cr_n = c\,\xi, \text{ by continuity of } f.$$

Hence, $f(x) = cx$, for all $x \in \mathbf{R}$.

Note: If $f : \mathbf{R} \to \mathbf{R}$ satisfies the functional equation (1), then the assumption that f is continuous at a *single point* implies that f is continuous *everywhere* on $\mathbf{R}$.

Now, for $a > 0$, the continuous function $f(x) = a^x$, $x \in \mathbf{R}$, satisfies the *functional equation.*

$$f(x+y) - f(x) f(y), \text{ for all } x,\ y \in \mathbf{R} \qquad ...(2)$$

1. *The only non-zero continuous function f on* $\mathbf{R}$ *that satisfies the functional equation* (2) *is the exponential function.*

   By hypothesis $f(x) \neq 0$, for any x. Moreover, from (2), we obtain

$$f(x) = [f(x/2)]^2 > 0, \text{ for all } x \in \mathbf{R}.$$

We may, therefore, take the logarithm of $f(x)$, (say, to the base e), and so taking $g(x) = \log_e f(x)$, $x \in \mathbf{R}$, it follows that $g(x+y) = g(x) + g(y)$, $\forall\ x, y \in \mathbf{R}$ and g, being composition of two continuous functions $f(x)$ and $\log_e x$, is itself continuous on $\mathbf{R}$.

Thus $g(x) = cx$, $\forall\ x \in \mathbf{R}$, c being constant.

Hence $f(x) = e^{cx} = (e^c)^x = a^x$, for all $x \in \mathbf{R}$.

The continuous function $f(x) = \log_a x$ $(a > 0, a \neq 1)$, $x > 0$, satisfies the *functional equation*,

$$f(xy) = f(x) + f(y), \text{ for all } x, y > 0 \qquad \ldots(3)$$

This relation, together with continuity, is enough to characterize the logarithmic function.

2. *The only non-zero function f continuous on* $\mathbf{R}^+$ *that satisfies the functional equation* (3) *is the logarithmic function.*

For $x > 0$, we put $x = e^t$, *i.e.*, $t = \log_e x$, where $t \in \mathbf{R}$.

Moreover, we define another function g on $\mathbf{R}$ by, $g(t) = f(e^t)$, $t \in \mathbf{R}$.

Then g is continuous on $\mathbf{R}$ and satisfies, for all $t, s \in \mathbf{R}$,

$$g(t+s) = f(e^{t+s}) = f(e^t e^s) = f(e^t) + f(e^s) = g(t) + g(s),$$

and so $g(t) = ct$, $t \in \mathbf{R}$, c being constant. Thus $f(x) = c \log e^x$, $x \in \mathbf{R}^+$. But c cannot be zero, since f is non-zero. Therefore, by taking $a = e^{1/c}$, we get

$$f(x) = \log_a x,\ x \in \mathbf{R}^+\ (a > 0, a \neq 1).$$

The proofs of the following are left as an exercise to the reader.

3. *The only non-zero continuous functions on* $\mathbf{R}$ *that satisfy the functional equation,*

$$f(x+y) + f(x-y) = 2f(x)f(y),\ \forall\ x, y \in \mathbf{R}$$

are the trigonometric and the hyperbolic cosines.

4. *The only non-zero continuous functions on* $\mathbf{R}$ *that satisfy the functional equation,*

$$f(x+y)f(x-y) = (f(x))^2 - (f(y))^2, \text{ for all } x, y \in \mathbf{R}$$

are the scalar ($f(x) = cx$), *the trigonometric and the hyperbolic sines.*

6. FUNCTIONS OF BOUNDED VARIATION

In this section we shall discuss the concept of the functions of bounded variation. The concept is closely associated to that of monotonic functions and has wide application in Mathematics. Presently we shall use these functions in Riemann-Stieltjes integrals and Fourier series.

A finite set P of points, $x_0, x_1, x_2, \ldots, x_n$, where

$$a = x_0 \leq x_1 \leq x_2 \leq \ldots \leq x_n = b$$

is called a *partition* of the interval $[a, b]$. Clearly any number of partition of $[a, b]$ are possible.

The intervals $[x_0, x_1]$, $[x_1, x_2]$, ..., $[x_{n-1}, x_n]$ are the sub-intervals of the partition. The ith sub-interval $[x_{i-1}, x_i]$, as also its length $x_i - x_{i-1}$, is denoted by Δx_i.

Let a function f be defined on an interval $[a, b]$ and $P = \{a = x_0, x_1, ..., x_n = b\}$ be a partition of $[a, b]$. Consider the sum $\sum_{i=1}^{n} |f(x_i) - f(x_{i-1})|$.

Since one such sum corresponds to each partition of $[a, b]$, the set of these sums is infinite. *If this set of sums is bounded above, the function f is said to be of bounded variation and the supremum of the set is called the total variation of f on* $[a, b]$, *and is denoted by the symbol,* $V(f, a, b)$. Thus

$$V(f, a, b) = \text{Sup} \sum \left| f(x_i) - f(x_{i-1}) \right|,$$

the supremum being taken over all partitions of $[a, b]$.

If there is no scope for confusion and the interval in question is clear from the context, we shall abbreviate the symbol to $V(f)$.

Thus the function f is said to be of bounded variation on $[a, b]$ if and only if its total variation is finite, *i.e.*,

$$V(f, a, b) < +\infty$$

Note: Since for $x \le c \le y$,

$$|f(y) - f(x)| \le |f(y) - f(c)| + |f(c) - f(x)|$$

the sum $\Sigma \left| f(x_i) - f(x_{i-1}) \right|$ cannot decrease (it can, in fact only increase) by the introduction of additional points to a partition of the interval.

6.1 Illustrative Examples

Example 1. A bounded monotonic function is a function of bounded variation.

- Let f be monotonic increasing on an interval $[a, b]$, and $P = \{a = x_0, x_1, ..., x_n = b\}$ be any partition of $[a, b]$. Then

$$\sum_{i=1}^{n} \left| f(x_i) - f(x_{i-1}) \right| = \sum_{i=1}^{n} \left\{ f(x_i) - f(x_{i-1}) \right\} = f(b) - f(a)$$

$$\therefore \quad V(f, a, b) = \text{Sup}\, \Sigma \left| f(x_i) - f(x_{i-1}) \right| = f(b) - f(a)$$

Thus, a monotone increasing bounded function is of bounded variation on $[a, b]$.

Similarly, it may be shown that a monotone decreasing bounded function is of bounded variation, with total variation $= f(a) - f(b)$.

Thus for a bounded monotonic f,

$$V(f) = |f(b) - f(a)|$$

Example 2. To show that a continuous function may not be a function of bounded variation, consider a function f, where

$$f(x) = \begin{cases} x \sin \pi/x, & \text{when } 0 < x \le 1 \\ 0, & \text{when } x = 0 \end{cases}$$

- Clearly f is continuous on $[0, 1]$.

Let us choose the partition

$$P = \left\{0, \frac{2}{2n+1}, \frac{2}{2n-1}, \ldots, \frac{2}{5}, \frac{2}{3}, 1\right\}$$

$$\therefore \sum_i \left|f(x_i) - f(x_{i-1})\right| = \left|f(1) - f\left(\frac{2}{3}\right)\right| + \left|f\left(\frac{2}{3}\right) - f\left(\frac{2}{5}\right)\right| + \ldots + \left|f\left(\frac{2}{2n+1}\right) - f(0)\right|$$

$$= \frac{2}{3} + \left(\frac{2}{3} + \frac{2}{5}\right) + \left(\frac{2}{5} + \frac{2}{7}\right) + \ldots + \left(\frac{2}{2n-1} + \frac{2}{2n+1}\right) + \frac{2}{2n+1}$$

$$= 4\left[\frac{1}{3} + \frac{1}{5} + \frac{1}{7} + \ldots + \frac{1}{2n+1}\right]$$

Since the infinite series $\frac{1}{3} + \frac{1}{5} + \frac{1}{7} + \ldots$ is divergent, its partial sums sequence $\{S_n\}$, where

$$S_n = \frac{1}{3} + \frac{1}{5} + \frac{1}{7} + \ldots + \frac{1}{2n+1}$$

is not bounded above.

Thus $\sum_i \left|f(x_i) - f(x_{i-1})\right|$ can be made arbitrarily large by taking n sufficiently large. Consequently, $V(f, 0, 1) \to \infty$ and so f is not of bounded variation on $[0, 1]$.

Remark: It may also be seen that a function of bounded variation is not necessarily continuous. The function $f(x) = [x]$, where $[x]$ denotes the greatest integer not greater than x, is a function of bounded variation on $[0, 2]$ but is not continuous.

Example 3. If the derivative f' exists and is bounded on $[a, b]$, then the function f is of bounded variation on $[a, b]$.

■ Since f' is bounded on $[a, b]$, therefore there exists K such that

$$|f'(x)| \le K, \ \forall\ x \in [a, b]$$

Let $P = \{a = x_0, x_1, \ldots, x_n = b\}$ be any partition of $[a, b]$.

$$\therefore \quad \sum_i \left|f(x) - f(x_{i-1})\right| = \sum_i (x_i - x_{i-1})\,|f'(\xi)|, \ \xi \in\,]a, b[$$

$$\le K(b - a)$$

$\Rightarrow$ $V(f, a, b)$ is finite and therefore f is of bounded variation.

Note: Boundedness of f' is a sufficient condition. It is not necessary.

Example 4. A function of bounded variation is necessarily bounded.

■ Let a function f be of bounded variation on $[a, b]$. For any $x \in [a, b]$, consider the partition $\{a, x, b\}$, consisting of just three points. Now

$$|f(x) - f(a)| + |f(b) - f(x)| \le V(f, a, b)$$

$$\Rightarrow \quad |f(x) - f(a)| \le V(f, a, b)$$

Again

$$|f(x)| = |f(a) + f(x) - f(a)|$$
$$\le |f(a)| + |f(x) - f(a)| \le |f(a)| + V(f, a, b)$$

$\Rightarrow$ f is bounded on $[a, b]$.

6.2 Some Properties of Functions of Bounded Variation

We shall see later (§ 6.3, Th. 3) that there is a close relation between functions of bounded variation and monotonic functions but there is one difference which is worthy of note. The functions of bounded variation are closed with respect to the arithmetic operations of addition and multiplication whereas the sum or the product of two monotonic functions need not be monotonic. For example, x and x^2 are monotonic in $[0, 1]$, but $x - x^2$ is not. Similarly, x is monotonic in $[-1, 1]$ but x^2 is not. In this section, we shall consider some properties of the functions of bounded variation, and in particular the arithmetic operations of addition and multiplication on them.

1. *The sum (difference) of two functions of bounded variation is also of bounded variation.*

Let f and g be two functions of bounded variation on $[a, b]$.

For any partition $P = \{a = x_0, x_1, \ldots, x_n = b\}$ of $[a, b]$, we have

$$\sum_i |(f+g)(x_i) - (f+g)(x_{i-1})| = \sum_i |\{f(x_i) + g(x_i)\} - \{f(x_{i-1}) + g(x_{i-1})\}|$$
$$\le \sum_i |f(x_i) - f(x_{i-1})| + \sum_i |g(x_i) - g(x_{i-1})|$$
$$\le V(f, a, b) + V(g, a, b)$$

$$\Rightarrow \quad V(f+g, a, b) \le V(f, a, b) + V(g, a, b)$$

and $(f + g)$ is of bounded variation.

Similarly, it may be shown that $(f - g)$ is of bounded variation and its total variation,

$$V(f - g) \le V(f) + V(g)$$

Corollary. *If f and g are monotonic increasing on* $[a, b]$, *then $f - g$ is of bounded variation on* $[a, b]$.

The converse of the corollary is also true (see § 6.3).

Note: If C is a constant, the sums $\Sigma |f(x_i) - f(x_{i-1})|$ and therefore the total variation function, $V(f)$ are same for f and $f - C$.

2. *The product of two functions of bounded variation is also of bounded variation.*

Let f and g be two functions of bounded variation on $[a, b]$.

Evidently f and g are bounded and therefore a number k exists such that

$$\left|f(x)\right| \le k, \left|g(x)\right| \le k, \ \forall \ x \in [a, b]$$

For any partition $\{a = x_0, x_1, \ldots, x_n = b\}$, we have

$$\sum_i \left|(fg)(x_i) - (fg)(x_{i-1})\right| = \sum_i \left|f(x_i)\, g(x_i) - f(x_{i-1}) g(x_{i-1})\right|$$

$$= \sum_i \left|f(x_i)\{g(x_i) - g(x_{i-1})\} + g(x_{i-1})\{f(x_i) - f(x_{i-1})\}\right|$$

$$\le \sum_i \left[\left|f(x_i)\right|\left|g(x_i) - g(x_{i-1})\right| + \left|g(x_{i-1})\right|\left|f(x_i) - f(x_{i-1})\right|\right]$$

$$\le k \sum_i \left|g(x_i) - g(x_{i-1})\right| + k \sum_i \left|f(x_i) - f(x_{i-1})\right|$$

$$\le kV(g) + kV(f)$$

$\Rightarrow$ the function (fg) is of bounded variation on $[a, b]$.

Note: Theorems like the above, could not be applied to quotients of functions because the reciprocal of a function of bounded variation need not be of bounded variation. For example, if $f(x) \to 0$ as $x \to x_0$, then $1/f$ will not be bounded and therefore cannot be of bounded variation on any interval containing x_0. Therefore, to consider quotient, we avoid functions whose values become arbitrarily close to zero.

3. *If f is a function of bounded variation on $[a, b]$ and if there exists a positive number k such that $\left|f(x)\right| \ge k$, for all $x \in [a, b]$, then $1/f$ is also of bounded variation on $[a, b]$.*

For any partition $\{a = x_0, x_1, \ldots, x_n = b\}$, we have

$$\sum_i \left|\left(\frac{1}{f}\right)(x_i) - \left(\frac{1}{f}\right)(x_{i-1})\right| = \sum_i \left|\frac{1}{f(x_i)} - \frac{1}{f(x_{i-1})}\right|$$

$$= \sum_i \left|\frac{f(x_{i-1}) - f(x_i)}{f(x_i)\, f(x_{i-1})}\right|$$

$$\le \frac{1}{k^2} \Sigma \left|f(x_i) - f(x_{i-1})\right|$$

$$\le V(f, a, b)/k^2$$

$\Rightarrow$ $1/f$ is of bounded variation on $[a, b]$.

4. *If f is of bounded variation on $[a, b]$ then it is also of bounded variation on $[a, c]$ and $[c, b]$, where c is a point of $[a, b]$, and conversely. Also*

$$V(f, a, b) = V(f, a, c) + V(f, c, b)$$

(*i*) Let, first, f be of bounded variation on $[a, b]$.

Let $P_1 = \{a = x_0, x_1, \ldots, x_m = c\}$

and

$$P_2 = \{c = y_0, y_1, ..., y_n = b\}$$

be any two partitions of $[a, c]$ and $[c, b]$ respectively. Evidently,

$$P = P_1 \cup P_2 = \{a = x_0, ..., x_m, y_0, ..., y_n = b\}$$

is a partition of $[a, b]$.

We have

$$\left\{\sum_{i=1}^{m} \left|f(x_i) - f(x_{i-1})\right| + \sum_{i=1}^{n} \left|f(y_i) - f(y_{i-1})\right|\right\} \le V(f, a, b)$$

$$\Rightarrow \quad \sum_{i=1}^{m} \left|f(x_i) - f(x_{i-1})\right| \le V(f, a, b)$$

and

$$\sum_{i=1}^{n} \left|f(y_i) - f(y_{i-1})\right| \le V(f, a, b)$$

$\Rightarrow$ f is of bounded variation on $[a, c]$ and $[c, b]$ both.

(*ii*) Let, now, f be of bounded variation on $[a, c]$ and $[c, b]$ both.

Let $P = \{a = z_0, z_1, z_2, ..., z_n = b\}$ be any partition of $[a, b]$. If it does not contain the point c, let us consider the partition $P^* = P \cup \{c\}$. Let $c \in \Delta z_r$, *i.e.*, $z_{r-1} \le c \le z_r$, $r < n$.

We have

$$\begin{aligned}\sum_{i=1}^{n} \left|f(z_i) - f(z_{i-1})\right| &= \sum_{i=1}^{r-1} \left|f(z_i) - f(z_{i-1})\right| + \left|f(z_r) - f(z_{r-1})\right| \\ &\quad + \sum_{i=r+1}^{n} \left|f(z_i) - f(z_{i-1})\right| \\ &\le \left\{\sum_{i=1}^{r-1} \left|f(z_i) - f(z_{i-1})\right| + \left|f(c) - f(z_{r-1})\right|\right\} \\ &\quad + \left\{\left|f(z_r) - f(c)\right| + \sum_{i=r+1}^{n} \left|f(z_i) - f(z_{i-1})\right|\right\} \\ &\le V(f, a, c) + V(f, c, b)\end{aligned}$$

$\Rightarrow$ f is of bounded variation on $[a, b]$ if it is of bounded variation on $[a, c]$ and $[c, b]$ both, and then

$$V(f, a, b) \le V(f, a, c) + V(f, c, b) \qquad ...(1)$$

(*iii*) Let $\varepsilon > 0$ be an arbitrary number.

Since $V(f, a, c)$ and $V(f, c, b)$ are the total variations of f on $[a, c]$ and $[c, b]$ respectively, there exist partitions

$$P_1 = \{a = x_0, x_1, x_2, ..., x_m = c\}$$
$$P_2 = \{c = y_0, y_1, y_2, ... y_n = b\}$$

of $[a, c]$ and $[c, b]$, respectively, such that

$$\sum_{i=1}^{m} \left| f(x_i) - f(x_{i-1}) \right| > V(f, a, c) - \frac{1}{2}\varepsilon \qquad ...(2)$$

$$\sum_{i=1}^{n} \left| f(y_i) - f(y_{i-1}) \right| > V(f, c, b) - \frac{1}{2}\varepsilon \qquad ...(3)$$

From (2) and (3), we have

$$\sum_{i=1}^{m} \left| f(x_i) - f(x_{i-1}) \right| + \sum_{i=1}^{n} \left| f(y_i) - f(y_{i-1}) \right| > V(f, a, c) + V(f, c, b) - \varepsilon$$

$$\Rightarrow V(f, a, b) \geq V(f, a, c) + V(f, c, b) - \varepsilon$$

But since ε is an arbitrary positive number, we get

$$V(f, a, b) \geq V(f, a, c) + V(f, c, b) \qquad ...(4)$$

Thus, from equation (1) and (4), we get

$$V(f, a, b) = V(f, a, c) + V(f, c, b)$$

6.3 Variation Function

Let f be a function of bounded variation on $[a, b]$ and x, a point of $[a, b]$. Then the total variation of f, $V(f, a, x)$ on $[a, x]$, which clearly is a function of x, is called the *total variation function* or simply the *variation function* of f and is denoted by $v_f(x)$, and when there is no scope for confusion, simply by $v(x)$. Thus,

$$v_f(x) = V(f, a, x), \ a \leq x \leq b$$

If x_1, x_2 are two points of $[a, b]$ such that $x_2 > x_1$, then

$$0 \leq \left| f(x_2) - f(x_1) \right| \leq V(f, x_1, x_2)$$
$$= V(f, a, x_2) - V(f, a, x_1) = v_f(x_2) - v_f(x_1)$$
$$\Rightarrow \qquad v_f(x_2) \geq v_f(x_1)$$

i.e., $v_f(x)$ is a monotone increasing function on $[a, b]$.

Remark: If f is of bounded variation on [a, b], then $v_f \pm f$ is a monotone increasing function on [a, b].

The following theorem characterizes the functions of bounded variation.

Theorem 3. ***Jordan Theorem.*** *A function of bounded variation is expressible as a difference of two monotone increasing functions.*

or

If f is a function of bounded variation on $[a, b]$, then there exist monotone increasing functions p and q on $[a, b]$ such that for $a \leq x \leq b$,

$$f(x) = p(x) - q(x) \quad ...(1)$$
$$v_f(x) = p(x) + q(x) \quad ...(2)$$

Let us define p and q by

$$p = \frac{1}{2}(v_f + f)$$
$$q = \frac{1}{2}(v_f - f)$$

so that equation (1) and (2) hold.

Now for $x_2 > x_1$, we have

$$\begin{aligned} p(x_2) - p(x_1) &= \frac{1}{2}[v_f(x_2) + f(x_2)] - \frac{1}{2}[v_f(x_1) + f(x_1)] \\ &= \frac{1}{2}[v_f(x_2) - v_f(x_1)] + \frac{1}{2}[f(x_2) - f(x_1)] \\ &= \frac{1}{2}[V(f, x_1, x_2) + \{f(x_2) - f(x_1)\}] \end{aligned}$$

But, since $V(f, x_1, x_2) \geq |f(x_2) - f(x_1)|$, we get

$$p(x_2) - p(x_1) \geq 0 \quad \Rightarrow \quad p(x_2) \geq p(x_1)$$

so that p is monotone increasing on $[a, b]$.

Again, we have

$$\begin{aligned} q(x_2) - q(x_1) &= \frac{1}{2}[v_f(x_2) - f(x_2)] - \frac{1}{2}[v_f(x_1) - f(x_1)] \\ &= \frac{1}{2}[V(f, x_1, x_2) - \{f(x_2) - f(x_1)\}] \geq 0 \end{aligned}$$

$$\Rightarrow \qquad q(x_2) \geq q(x_1)$$

so that q is monotone increasing on $[a, b]$.

Hence the result.

With the help of the result proved above and that of § 6.1, Example 1 *and* § 6.2, we may state that *a function is of bounded variation on an interval if and only if it can be expressed as the difference of two monotone increasing functions.*

Note: The functions p and q are respectively called the *positive* and the *negative* variation function of f.
We shall now prove that the variation function of a continuous function of bounded variation is itself continuous, and conversely.

Theorem 4. *Variation function of a continuous function.* *The variation function of a function f of bounded variation is continuous if and only if f is a continuous function.*

Necessary. Let the variation function $v(x)$ of f be continuous at a point c of $[a, b]$.

Let $\varepsilon > 0$ be an arbitrary number.

Because of continuity of $v(x)$ at c, $\delta > 0$ exists such that

$$\left|v(x)-v(c)\right|<\varepsilon, \text{ for } |x-c|<\delta \quad ...(1)$$

Also

$$\left|f(x)-f(c)\right| \le v(x)-v(c) \quad \text{if } x>c \quad ...(2)$$

$$\left|f(x)-f(c)\right| \le v(c)-v(x) \quad \text{if } x<c \quad ...(3)$$

Hence, from equation (1), (2) and (3), we get

$$\left|f(x)-f(c)\right|<\varepsilon, \text{ for } |x-c|<\delta$$

which implies that f is continuous at c.

Sufficient. Let now f be continuous at c, so that for an arbitrary $\varepsilon > 0$ there exists $\delta > 0$ such that

$$\left|f(x)-f(c)\right|<\frac{1}{2}\varepsilon, \text{ for } |x-c|<\delta$$

Again since $V(f, c, b)$ is total variation of f on $[c, b]$, there exists a partition $P = \{c = x_0, x_1, ..., x_n = b\}$ of $[c, b]$ such that

$$\sum_{i=1}^{n}\left|f(x_i)-f(x_{i-1})\right|>V(f,c,b)-\frac{1}{2}\varepsilon \quad ...(4)$$

Let us assume that the length of the first sub-interval $x_1 - c$ is less than δ, for, otherwise we make it so, by introducing additional points to the partition P. Also, (§ 6, note) the sum $\sum_i \left|f(x)-f(x_{i-1})\right|$ cannot decrease by the introduction of more points to P, so that the inequality (4) remains unaffected.

Thus, let $0 < x_1 - c < \delta$ so that

$$\left|f(x_1)-f(c)\right|<\frac{1}{2}\varepsilon \quad ...(5)$$

Again (4) gives on using (5),

$$V(f,c,b)-\frac{1}{2}\varepsilon<\frac{1}{2}\varepsilon+\sum_{i=2}^{n}\left|f(x_i)-f(x_{i-1})\right| \le \frac{1}{2}\varepsilon+V(f,x_1,b)$$

$$\Rightarrow \quad V(f,c,b)-V(f,x_1,b)<\varepsilon$$

or

$$V(f,c,x_1)<\varepsilon$$

$$\therefore \quad v(x_1)-v(c)=V(f,a,x_1)-V(f,a,c)=V(f,c,x_1)<\varepsilon$$

Thus

$$-\varepsilon<0<v(x_1)-v(c)<\varepsilon, \text{ when } 0<x_1-c<\delta$$

$$\Rightarrow \quad \lim_{x\to c+} v(x)=v(c)$$

Similarly, it can be shown that

$$\lim_{x\to c-} v(x)=v(c)$$

Hence, $v(x)$ is continuous at c.

Again, since c is any point of $[a, b]$, we deduce that $v_f(x)$ is continuous on $[a, b]$ if and only if f is continuous on $[a, b]$.

Corollary. Continuity of f implies the continuity of $v_f(x)$ and therefore of the positive and negative variation functions p and q conversely.

Thus, *a continuous function is of bounded variation if and only if it can be expressed as a difference of two continuous monotonic increasing functions.*

Example 5. Determine whether or not the following function f is of bounded variation on $[0, 1]$.

$$f(x) = x^2 \sin(1/x), \text{ if } x \neq 0, \ f(0) = 0$$

■ Clearly, f is continuous on $[0, 1]$, and

$$f'(x) = 2x \sin(1/x) - \cos(1/x), \text{ if } x \neq 0$$

$$f'(0) = \lim_{x \to 0} \frac{x^2 \sin(1/x) - 0}{x} = 0$$

Also

$$|f'(x)| \leq 3, \text{ for all } x \in [0, 1]$$

Thus we find that f' exists and is bounded on $[0, 1]$. So that (by the hypothesis of example 3, § 6.1) the function f is of bounded variation on $[0, 1]$.

Example 6. f is a function of bounded variation on $[a, b]$, and p and q are its positive and negative variation functions. If p_1, q_1 are two monotone increasing functions on $[a, b]$ such that $f = p_1 - q_1$, then show that $V(p) \leq V(p_1)$ and $V(q) \leq V(q_1)$, where V denotes total variation on $[a, b]$.

■ With usual notation, we have

$$p = \frac{1}{2}(v_f + f), \text{ and } q = \frac{1}{2}(v_f - f)$$

Given that $f = p_1 - q_1$, therefore for $x \in [a, b]$, we have

$$2p(x) = V(f, a, x) + p_1(x) - q_1(x)$$

$$\therefore \quad 2p(b) - 2p(a) = V(f, a, b) + p_1(b) - p_1(a) - \{q_1(b) - q_1(a)\}$$

i.e.,

$$2V(p) = V(f) + V(p_1) - V(q_1) \qquad ...(1)$$

But

$$V(f) = V(p_1 - q_1) \leq V(p_1) + V(q_1) \qquad ...(2)$$

From (1) and (2), it follows that

$$2V(p) \leq 2V(p_1), \text{ and so } V(p) \leq V(p_1)$$

Similarly, $V(q) \leq V(q_1)$.

Example 7. Compute the positive, negative and the total variation functions of

$$f(x) = 3x^2 - 2x^3, \text{ for } -2 \leq x \leq 2$$

■ Here

$$f'(x) = 6x(1 - x)$$

which vanishes for $x = 0, 1$ and is negative for $-2 \leq x < 0$ or $1 < x \leq 2$, and positive for $0 < x < 1$.

Hence, for $-2 \leq x \leq 0$, (when f is monotone decreasing)

$$v_f(x) = V(f, -2, x) = f(-2) - f(x)$$
$$= 28 - 3x^2 + 2x^3$$

$\therefore$
$$V(f, -2, 0) = 28$$

For $0 \le x \le 1$ (when f is monotone increasing)

$$V(f, 0, x) = f(x) - f(0) = 3x^2 - 2x^3$$

$\therefore$
$$v_f(x) = V(f, -2, x) = V(f, -2, 0) + V(f, 0, x)$$
$$= 28 + 3x^2 - 2x^3$$

and

$$V(f, -2, 1) = 29$$

For $1 \le x \le 2$ (when f is monotone decreasing)

$$V(f, 1, x) = f(1) - f(x) = 1 - 3x^2 + 2x^3$$

$\therefore$
$$v_f(x) = V(f, -2, x) = V(f, -2, 1) + V(f, 1, x)$$
$$= 30 - 3x^2 + 2x^3$$

Thus the total variation function on $-2 \le x \le 2$ is defined as:

$$v_f(x) = \begin{cases} 28 - 3x^2 + 2x^3, & \text{for } -2 \le x \le 0 \\ 28 + 3x^2 - 2x^3, & \text{for } \ 0 \le x \le 1 \\ 30 - 3x^2 + 2x^3, & \text{for } \ 1 \le x \le 2 \end{cases}$$

Since the positive variation function p is defined as:

$$p(x) = \frac{1}{2}\{v_f(x) + f(x)\}$$

$\therefore$
$$p(x) = \begin{cases} 14, & \text{for } -2 \le x \le 0 \\ 14 + 3x^2 - 2x^3, & \text{for } \ 0 \le x \le 1 \\ 15, & \text{for } \ 1 \le x \le 2 \end{cases}$$

Similarly, the negative variation function,

$$q(x) = \begin{cases} 14 - 3x^2 + 2x^3, & \text{for } -2 \le x \le 0 \\ 14, & \text{for } \ 0 \le x \le 1 \\ 15 - 3x^2 + 2x^3, & \text{for } \ 1 \le x \le 2 \end{cases}$$

Example 8. Compute the positive, negative and the total variation functions of f, where

$$f(x) = [x] - x \ (0 \le x \le 2)$$

■ The function f is monotone decreasing from 0 to 1 and from 1 to 2 and has discontinuities at 1 and 2. It may be restated as

$$f(x) = \begin{cases} -x, & \text{for } 0 \le x < 1 \\ 1 - x, & \text{for } 1 < x < 2 \\ 0, & \text{for } x = 1, 2 \end{cases}$$

For $0 \le x < 1$,

$$v_f(x) = V(f, 0, x) = f(0) - f(x) = x \equiv [x] + x$$

$$\begin{aligned} V(f, 0, 1) &= \underset{0<x<1}{\text{l.u.b.}} \left[V(f, 0, x) + V(f, x, 1)\right] \\ &= \underset{0<x<1}{\text{l.u.b.}} \left[x + |f(1) - f(x)|\right] \\ &= \underset{0<x<1}{\text{l.u.b.}} \left[x + |0 - x|\right] = 2 \end{aligned}$$

$\therefore$ $$\left[v_f(x)\right]_{x=1} = V(f, 0, 1) = 2$$

For $1 \le x < 2$,

$$V(f, 1, x) = f(1) - f(x) = -1 + x$$

$\therefore$ $$\begin{aligned} v_f(x) &= V(f, 0, x) = V(f, 0, 1) + V(f, 1, x) \\ &= 1 + x \equiv [x] + x \end{aligned}$$

$$\begin{aligned} V(f, 1, 2) &= \underset{1<x<2}{\text{l.u.b.}} \left[V(f, 1, x) + V(f, x, 2)\right] \\ &= \underset{1<x<2}{\text{l.u.b.}} \left[-1 + x + |f(2) - f(x)|\right] \\ &= \underset{1<x<2}{\text{l.u.b.}} \left[-1 + x + |0 - (1 - x)|\right] = 2 \end{aligned}$$

$\therefore$ $$\left[v_f(x)\right]_{x=2} = V(f, 0, 2) = V(f, 0, 1) + V(f, 1, 2) = 4$$

Thus the total variation function is defined on [0, 2] as

$$v_f(x) = [x] + x, \text{ for } 0 \le x \le 2$$

Accordingly the positive and negative variation functions, p and q, are defined as:

$$p(x) = [x], \text{ for } 0 \le x \le 2$$

$$q(x) = x, \text{ for } 0 \le x \le 2$$

Ex. 1. Show that $\sin x$ and $\cos x$ are of bounded variation over a finite interval.

Ex. 2. Show that a polynomial function is of bounded variation over any finite interval. Describe a method for finding the total variation of f on $[a, b]$, if the zeros of the derivative f' are known.

Ex. 3. Determine whether or not f is of bounded variation on [0, 1], where,

(*i*) $f(x) = \sqrt{x} \sin(1/x)$, if $x \ne 0$, $f(0) = 0$, and

(*ii*) $f(x) = x^2 \sin(1/x^2)$, if $x \ne 0$, $f(0) = 0$.

7. VECTOR-VALUED FUNCTIONS

7.1 The Euclidean Space $\mathbf{R}^n$

The set of all ordered n-tuples of real numbers is denoted by $\mathbf{R}^n$. Thus

$$\mathbf{R}^n = \{(t_1, t_2, ..., t_n) : t_j \in \mathbf{R} \text{ for } j = 1, 2, ..., n\}$$

The n-tuples $(t_1, t_2, ..., t_n)$, where $t_1, t_2, ..., t_n$ are real numbers, are *members* or *points* of $\mathbf{R}^n$, and $t_1, t_2, ..., t_n$ the *components* or the *coordinates* of the points of $\mathbf{R}^n$. Members of $\mathbf{R}^n$ will be denoted by bold faced letters **x**, **y**, **z**, .., so that each of **x**, **y**, **z**, etc. stands for an ordered n-tuple of real numbers.

0 denotes the point (0, 0, ..., 0).

If $\mathbf{x} = (x_1, x_2, ..., x_n)$ and $\mathbf{y} = (y_1, y_2, ..., y_n)$ be two points of $\mathbf{R}^n$, then we define *addition* and *scalar multiplication* by

(*i*) $\mathbf{x} + \mathbf{y} = (x_1 + y_1, x_2 + y_2, ..., x_n + y_n)$,

(*ii*) $c\mathbf{x} = (cx_1, cx_2, ..., cx_n)$, $c \in \mathbf{R}$.

(*iii*) Distance between two points **x** and **y**, denoted symbolically by $d(\mathbf{x}, \mathbf{y})$, is defined as

$$d(\mathbf{x}, \mathbf{y}) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + ... + (x_n - y_n)^2}$$

$$= \sqrt{\sum_{j=1}^{n} (x_j - y_j)^2}$$

If we define a non-negative number $|\mathbf{x}|$, where

$$|\mathbf{x}| = \sqrt{x_1^2 + x_2^2 + ... + x_n^2} = \sqrt{\sum_{j=1}^{n} x_j^2}$$

then $|\mathbf{x}|$ denotes the distance between **x** and 0, and is called the Norm of **x**. Also

$$d(\mathbf{x}, \mathbf{y}) = |\mathbf{x} - \mathbf{y}|$$

The set $\mathbf{R}^n$ equipped with the properties (*i*), (*ii*) and (*iii*), mentioned above, is called the **Euclidean space $\mathbf{R}^n$** of n dimension.

Cauchy-Schwarz Inequality

If $\mathbf{x} = (x_1, x_2, ..., x_n)$ and $\mathbf{y} = (y_1, y_2, ..., y_n)$ be two points of $\mathbf{R}^n$, then

$$\left| \sum_{i=1}^{n} x_i y_i \right| \le \left(\sum_{i=1}^{n} x_i^2 \right)^{1/2} \left(\sum_{i=1}^{n} y_i^2 \right)^{1/2}$$

We note that

$$0 \le \sum_{i=1}^{n} \sum_{j=1}^{n} \left(x_i y_j - x_j y_i \right)^2$$

$$= \sum_{i=1}^{n} x_i^2 \sum_{j=1}^{n} y_j^2 - 2 \sum_{i=1}^{n} x_i y_i \sum_{j=1}^{n} y_j x_j + \sum_{j=1}^{n} x_j^2 \sum_{i=1}^{n} y_i^2$$

$$= 2\sum_{i=1}^{n} x_i^2 \sum_{i=1}^{n} y_i^2 - 2\left(\sum_{i=1}^{n} x_i y_i \right)^2$$

It follows that

$$\left(\sum_{i=1}^{n} x_i y_i \right)^2 \le \sum_{i=1}^{n} x_i^2 \sum_{i=1}^{n} y_i^2$$

Taking positive square root of both sides, the Cauchy-Schwarz inequality is established.

Ex. For $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbf{R}^n$, show that

(*i*) $d(\mathbf{x}, \mathbf{y}) = d(\mathbf{y}, \mathbf{x}) \geq 0$,

(*ii*) $d(\mathbf{x}, \mathbf{y}) = 0$ iff $\mathbf{x} = \mathbf{y}$,

(*iii*) $d(\mathbf{x}, \mathbf{y}) \leq d(\mathbf{x}, \mathbf{z}) + d(\mathbf{z}, \mathbf{y})$ (Triangle inequality).

7.2 Vector-Valued Functions

If $f_1, f_2, ..., f_m$ be m real-valued functions on an interval $[a, b]$, then the corresponding mapping $\mathbf{f} = (f_1, f_2, ..., f_m)$ of $[a, b]$ into $\mathbf{R}^m$ is called a real vector valued function* of $\mathbf{R}$ into $\mathbf{R}^m$.

The domain of such a function is a subset of $\mathbf{R}$ and range, a subset of the m-dimensional space $\mathbf{R}^m$. The functions $f_1, f_2, ..., f_m$ are called its components.

Vector valued functions with domain $\mathbf{R}^k$ and range, $\mathbf{R}^n$ are also defined in a similar way.

The vector valued functions satisfy the following relations.

If $\mathbf{f} = (f_1, f_2, ..., f_m)$, $\mathbf{g} = (g_1, g_2, ..., g_m)$ be two vector valued functions, then

1. $\mathbf{f} = \mathbf{g}$ if and only if $f_j = g_j$, for $j = 1, 2, ..., m$
2. $\mathbf{f} + \mathbf{g} = (f_1 + g_1, f_2 + g_2, ..., f_m + g_m)$
3. $a\mathbf{f} = (af_1, af_2, ..., af_m)$, a any real number
   $0\mathbf{f} = (0, 0, ..., 0) = \mathbf{0}$
4. $\mathbf{f}$ is said to be continuous if and only if each f_j is continuous
5. $\mathbf{f}$ is differentiable when each f_j is differentiable, and then

$$\mathbf{f}' = (f_1', f_2', ..., f_m')$$

7.3 Vector-Valued Functions of Bounded Variation

Let $\mathbf{f}$ be a vector-valued function with domain $[a, b]$ and range $\mathbf{R}^m$. If $P = \{a = x_0, x_1, ..., x_n = b\}$ is any partition of $[a, b]$ and $\Delta\mathbf{f}_i = \mathbf{f}(x_i) - \mathbf{f}(x_{i-1})$ denotes the oscillation of $\mathbf{f}$ on Δx_i, then the *total variation* of $\mathbf{f}$ on $[a, b]$ is defined as

$$V(\mathbf{f}, a, b) = \text{Sup} \sum_{i=1}^{n} |\Delta\mathbf{f}_i| = \text{Sup} \sum_{i=1}^{n} \left|\mathbf{f}(x_i) - \mathbf{f}(x_{i-1})\right|,$$

the supremum being taken over all partitions of $[a, b]$.

$\mathbf{f}$ is said to be of *bounded variation* on $[a, b]$ if and only if $V(\mathbf{f}, a, b) < +\infty$, *i.e.*, $V(\mathbf{f}, a, b)$ is finite.

The function

$$v_{\mathbf{f}}(x) = V(\mathbf{f}, a, x),\ a \leq x \leq b$$

is called the *total variation function* of $\mathbf{f}$.

---

* The reader is already familiar with the vector functions in space (3-dimensional), where a vector function P can be expressed as $P = iP_1 + jP_2 + kP_3$ or equivalently as an ordered triplet (P_1, P_2, P_3). The definition given here is a generalisation of the same.

CHAPTER 8

Note: $V(\mathbf{f}, a, b)$ or $v_{\mathbf{f}}(x)$ are non-negative and can be zero only when $\mathbf{f}(x_i) = \mathbf{f}(x_{i-1})$ for all i, that is only when $\mathbf{f}$ is a constant function.

We shall now prove some theorems for the vector valued functions of bounded variation. As many properties of such functions can be reduced to that of the real-valued functions, we shall give the outline of the proofs, details may be provided by the reader himself.

1. $\mathbf{f}$ *is of bounded variation on* $[a, b]$ *if and only if each component function* f_j *is of bounded variation on* $[a, b]$.

For any partition $\{a = x_0, x_1, ..., x_n = b\}$ of $[a, b]$.

$$\left|f_j(x_i) - f_j(x_{i-1})\right| \le \left|\mathbf{f}(x_i) - \mathbf{f}(x_{i-1})\right| \le \sum_{j=1}^{m} \left|f_j(x_i) - f_j(x_{i-1})\right|$$

If we add these inequalities for $i = 1, 2, ..., n$ and take the least upper bound over all partitions of $[a, b]$, we get

$$V(f_j, a, b) \le V(f, a, b) \le \sum_{j=1}^{m} V(f_j, a, b)$$

which proves the theorem.

Corollary. Since each real-valued function of bounded variation is necessarily bounded, it follows that each function f_j and consequently $\mathbf{f}$ is bounded. Hence *a vector valued function of bounded variation is necessarily bounded.*

2. *If* $\mathbf{f}'$ *exists and is bounded on* $[a, b]$ *then* $\mathbf{f}$ *is of bounded variation on* $[a, b]$.

Since $\mathbf{f}$ is derivable, therefore each f_j is derivable.

Also, if $\left|\mathbf{f}'(x)\right| \le M$, for all $x \in [a, b]$, then

$$\left|f_j'(x)\right| \le \left|\mathbf{f}'(x)\right| \le M$$

Hence for any partition $\{a = x_0, x_1, ..., x_n = b\}$, we have by Lagrange's Mean Value Theorem,

$$\left|f_j(x_i) - f_j(x_{i-1})\right| = \left|x_i - x_{i-1}\right| \left|f_j'(\xi)\right|, \; x_{i-1} < \xi < x_i \le M(x_i - x_{i-1})$$

$$\therefore \quad \sum_{i=1}^{n} \left|f_j(x_i) - f_j(x_{i-1})\right| \le M \sum_{i=1}^{n} (x_i - x_{i-1}) = M(b - a)$$

Taking the supremum over all partition of $[a, b]$, we get

$$V(f_j, a, b) \le M(b - a)$$

so that each function f_j is of bounded variation on $[a, b]$. Consequently (by 1) $\mathbf{f}$ is of bounded variation on $[a, b]$.

3. *If* $\mathbf{f}$ *and* $\mathbf{g}$ *are of bounded variation then* $\mathbf{f} \pm \mathbf{g}$ *is also of bounded variation, and*

$$V(\mathbf{f} \pm \mathbf{g}, a, b) \le V(\mathbf{f}, a, b) + V(\mathbf{g}, a, b)$$

Let $\mathbf{f} = (f_1, f_2, ..., f_m)$, $\mathbf{g} = (g_1, g_2, ..., g_m)$

Since $\mathbf{f}$ and $\mathbf{g}$ are both of bounded variation, therefore each f_j and g_j is of bounded variation over $[a, b]$.

$\Rightarrow$ each $f_j \pm g_j$ is of bounded variation on $[a, b]$

$\Rightarrow$ $\mathbf{f} \pm \mathbf{g}$ is of bounded variation on $[a, b]$

For any partition $\{x_0, x_1, x_2, ..., x_n\}$ of $[a, b]$, we have

$$\sum_{i=1}^{n} \left|(\mathbf{f} \pm \mathbf{g})(x_i) - (\mathbf{f} \pm \mathbf{g})(x_{i-1})\right|$$

$$= \sum_{i=1}^{n} \left|\{\mathbf{f}(x_i) - \mathbf{f}(x_{i-1})\} \pm \{\mathbf{g}(x_i) - \mathbf{g}(x_{i-1})\}\right|$$

$$\le \sum_{i=1}^{n} \left|\mathbf{f}(x_i) - \mathbf{f}(x_{i-1})\right| + \sum_{i=1}^{n} \left|\mathbf{g}(x_i) - \mathbf{g}(x_{i-1})\right|$$

$$\le V(\mathbf{f}, a, b) + V(\mathbf{g}, a, b)$$

$$\Rightarrow \quad V(\mathbf{f} \pm \mathbf{g}, a, b) \le V(\mathbf{f}, a, b) + V(\mathbf{g}, a, b)$$

4. *For $a \le x \le y \le b$ prove that*

$$V(\mathbf{f}, a, y) \le V(\mathbf{f}, a, x) + V(\mathbf{f}, x, y) \qquad ...(1)$$

If $x = a$ or $y = x$, the result is trivial, for, $V(\mathbf{f}, x, x) = 0$.

Let $a < x < y \le b$.

Let $\varepsilon > 0$ be given.

There exists a partition $P = \{a = x_0, x_1, ..., x_n = y\}$ of $[a, y]$ such that

$$V(\mathbf{f}, a, y) - \varepsilon \le \sum_{P} \left|\mathbf{f}(x_i) - \mathbf{f}(x_{i-1})\right| \le V(\mathbf{f}, a, y) \qquad ...(2)$$

where $\sum_{P}$ denotes the summation over all points of P.

If x is not a member of P, we consider the partition $P^* = P \cup \{x\}$, for which (2) still holds.

Evidently P^* gives rise to the partitions P_1 and P_2 of $[a, x]$ and $[x, y]$ respectively such that $P^* = P_1 \cup P_2$.

Now from (2),

$$\sum_{P^*} \left|\mathbf{f}(x_i) - \mathbf{f}(x_{i-1})\right| \le V(\mathbf{f}, a, y)$$

or

$$\sum_{P_1} \left|\mathbf{f}(x_i) - \mathbf{f}(x_{i-1})\right| + \sum_{P_2} \left|\mathbf{f}(x_i) - \mathbf{f}(x_{i-1})\right| \le V(\mathbf{f}, a, y)$$

$$\Rightarrow \quad V(\mathbf{f}, a, x) + V(\mathbf{f}, x, y) \le V(\mathbf{f}, a, y) \qquad ...(3)$$

From (2) and (3), for any $\varepsilon > 0$, we get

$$V(\mathbf{f}, a, y) - \varepsilon \le V(\mathbf{f}, a, x) + V(\mathbf{f}, x, y) \le V(\mathbf{f}, a, y)$$

and since ε is arbitrary, result (1) follows.

5. *The total variation function $v_{\mathbf{f}}(x)$ is monotone increasing.*

For $a \le x_1 \le x_2 \le b$, we have

$$\begin{aligned} 0 &\le V(\mathbf{f}, x_1, x_2) \\ &= V(\mathbf{f}, a, x_2) - V(\mathbf{f}, a, x_1) \\ &= v_{\mathbf{f}}(x_2) - v_{\mathbf{f}}(x_1) \end{aligned}$$

$$\Rightarrow \qquad v_{\mathbf{f}}(x_2) \ge v_{\mathbf{f}}(x_1)$$

Corollary. *$v_{\mathbf{f}}(x)$ is a strictly monotone increasing function over $[x_1, x_2]$, unless* **f** *is constant on $[x_1, x_2]$.*

$$\because \qquad v_{\mathbf{f}}(x_2) = v_{\mathbf{f}}(x_1) \quad \Rightarrow \quad V(\mathbf{f}, x_1, x_2) = 0$$

$$\therefore \qquad 0 \le V(f_j, x_1, x_2) \le V(\mathbf{f}, x_1, x_2) = 0, \text{ for } j = 1, 2, \ldots, m$$

$$\Rightarrow \qquad V(f_j, x_1, x_2) = 0, \text{ for all } j = 1, 2, \ldots, m$$

$\Rightarrow$ all the components $f_1, f_2, \ldots, f_m$ are constant over $[x_1, x_2]$

$\Rightarrow$ **f** is constant over $[x_1, x_2]$

6. *If a vector valued function* **f**, *with domain $[a, b]$ and range a subset of $\mathbf{R}^m$, is of bounded variation on $[a, b]$, then $v_{\mathbf{f}}(x)$ is continuous at $c \in [a, b]$ if and only if* **f** *is continuous at c.*

(*i*) Let, *first*, **f** be continuous at $c \in [a, b]$.

Consequently each f_j and therefore each $v_{f_j}(x)$ is continuous at c.

Thus for $\varepsilon > 0$, $\delta > 0$ can be found such that for $j = 1, 2, \ldots, m$

$$\left| V_{f_j}(c+h) - v_{f_j}(c) \right| < \varepsilon/m, \text{ when } |h| \le \delta$$

Now, for $h > 0$,

$$\begin{aligned} 0 \le v_{\mathbf{f}}(c+h) - v_{\mathbf{f}}(c) &= V(\mathbf{f}, c, c+h) \\ &= \sum_{j=1}^{m} V(f_j, c, c+h) \\ &= \sum_{j=1}^{m} \left\{ v_{f_j}(c+h) - v_{f_j}(c) \right\} \end{aligned} \qquad \ldots(1)$$

Similarly for $h < 0$,

$$0 \le v_{\mathbf{f}}(c) - v_{\mathbf{f}}(c+h) \le \sum_{j=1}^{m} \left\{ v_{f_j}(c) - v_{f_j}(c+h) \right\} \qquad \ldots(2)$$

so that from (1) and (2), for $|h| \le \delta$, we have

$$\left| v_{\mathbf{f}}(c+h) - v_{\mathbf{f}}(c) \right| \le \sum_{j=1}^{m} \left| v_{f_j}(c+h) - v_{f_j}(c) \right| \le \frac{m \cdot \varepsilon}{m} = \varepsilon$$

$$\Rightarrow \qquad v_{\mathbf{f}}(x) \text{ is continuous at } c.$$

(*ii*) Let, now, $v_{\mathbf{f}}(x)$ be continuous at $c \in [a, b]$, so that for $\varepsilon > 0, \exists\, \delta > 0$ such that

$$\left|v_{\mathbf{f}}(x) - v_{\mathbf{f}}(c)\right| < \varepsilon, \text{ for } |x - c| < \delta \qquad \text{...(3)}$$

Again

$$\begin{aligned}\left|\mathbf{f}(x) - \mathbf{f}(c)\right| &\le V(\mathbf{f}, c, x), \text{ if } x > c \\ &= V(\mathbf{f}, a, x) - V(\mathbf{f}, a, c) \\ &= v_{\mathbf{f}}(x) - v_{\mathbf{f}}(c), \text{ if } x > c \qquad \text{...(4)}\end{aligned}$$

and

$$\left|\mathbf{f}(c) - \mathbf{f}(x)\right| \le v_{\mathbf{f}}(c) - v_{\mathbf{f}}(x), \text{ if } x < c \qquad \text{...(5)}$$

Hence for (3), (4) and (5), we get

$$\left|\mathbf{f}(x) - \mathbf{f}(c)\right| < \varepsilon, \text{ for } |x - c| < \delta$$

$\Rightarrow$ **f** is continuous at c.

Integration of Vector-Value Functions

7.4* If α is a monotonic increasing function on $[a, b]$, then $\mathbf{f} \in R(\alpha)$ means that $f_j \in R(\alpha)$, for $j = 1, 2, ..., m$ and we define

$$\int_a^b \mathbf{f}\, d\alpha = \left(\int_a^b f_1\, d\alpha, \int_a^b f_2\, d\alpha, ..., \int_a^b f_m\, d\alpha\right)$$

Thus $\int_a^b \mathbf{f}\, d\alpha$ is a point in $\mathbf{R}^m$ whose jth component is $\int_a^b f_j\, d\alpha$.

The following results are easy consequences of the definition:

(*a*) $\int_a^b (\mathbf{f} + \mathbf{g})\, d\alpha = \int_a^b \mathbf{f}\, d\alpha + \int_a^b \mathbf{g}\, d\alpha$

(*b*) For $a < c < b$, $\mathbf{f} \in R(\alpha)$ on $[a, b]$ if and only if $\mathbf{f} \in R(\alpha)$ on $[a, c]$ and $[c, b]$ both and then

$$\int_a^b \mathbf{f}\, d\alpha = \int_a^c \mathbf{f}\, d\alpha + \int_c^b \mathbf{f}\, d\alpha$$

(*c*) If $\mathbf{f} \in R(\alpha)$, then $\mathbf{f}^2 \in R(\alpha)$, $|\mathbf{f}| \in R(\alpha)$, and

$$\left|\int_a^b \mathbf{f}\, d\alpha\right| \le \int_a^b |\mathbf{f}|\, d\alpha.$$

(*d*) $\mathbf{f} \in R(\alpha_1)$ and $\mathbf{f} \in R(\alpha_2)$, then $\mathbf{f} \in R(\alpha_1 + \alpha_2)$, and

* See, The Riemann-Stieltjes Integral.

CHAPTER 8

$$\int_a^b \mathbf{f}\, d(\alpha_1 + \alpha_2) = \int_a^b \mathbf{f}\, d\alpha_1 + \int_a^b \mathbf{f}\, d\alpha_2$$

(*e*) If **f** is continuous then $\mathbf{f} \in R(\alpha)$.

(*f*) If $\mathbf{f} \in R$ on $[a, b]$ and $a \le x \le b$, we define

$$\mathbf{F}(x) = \int_a^x \mathbf{f}(t)\, dt$$

then **F** is continuous on $[a, b]$. Further if **f** is continuous at a point x_0 of $[a, b]$ then **F** is differentiable at x_0, and

$$\mathbf{F}'(x_0) = \mathbf{f}(x_0).$$

(*g*) If $\mathbf{f} \in R$ on $[a, b]$, and if there is a differentiable function **F** on $[a, b]$ such that $\mathbf{F}' = \mathbf{f}$, then

$$\int_a^b \mathbf{f}\, dx = \mathbf{F}(b) = \mathbf{F}(a).$$

(*h*) If $\mathbf{f} \in R$ and $\alpha' \in R$ on $[a, b]$, then $\mathbf{f} \in R(\alpha)$, and

$$\int_a^b \mathbf{f}\, d\alpha = \int_a^b \mathbf{f}(x)\alpha'(x)\, dx.$$

9 The Riemann Integral

(Integration of Bounded Functions on R)

At elementary stage, the subject of integration is generally introduced as the inverse of differentiation, so that a function F is called an integral of a given function f, if $F'(x) = f(x)$, for all values of x belonging to the domain of the function f. Historically speaking the subject arose in connection with the evaluation of areas of plane regions and thus amounted to finding out the limit of a sum when the number of terms tended to infinity, each term tending to zero. Realisation that the subject could be looked upon as inverse of differentiation came afterwards. The reference to integration from summation point of view was always associated with the geometrical concepts.

To formulate an independent theory of integration, the German mathematician, Riemann, gave a purely arithmetic treatment to the subject and thus developed the subject entirely free from the intuitive dependence on geometrical concepts. Many refinements and generalisations of the subject followed, the most noteworthy being *Lebesgue theory* of integration.

The present chapter is, thus, based on a definition of the Riemann integral which depends very explicitly on the order structure of the real line. Accordingly, we begin by discussing the integration of real bounded functions on intervals. Integration over sets other than intervals is beyond the scope of the present discussion.

The function will always be assumed to be bounded unless otherwise stated.

1. DEFINITIONS AND EXISTENCE OF THE INTEGRAL

Let $[a, b]$ be a given closed interval.

Partition. By a partition of $[a, b]$, we mean a finite set P of points $x_0, x_1, x_2, ..., x_n$,

where

$$a = x_0 \leq x_1 \leq ... \leq x_{n-1} \leq x_n = b$$

The partition P consists of $n + 1$ points. Clearly any number of partitions of $[a, b]$ can be considered.

$[x_0, x_1], [x_1, x_2], ..., [x_{i-1}, x_i], ..., [x_{n-1}, x_n]$ are the sub-intervals of $[a, b]$. We shall use the same symbol Δx_i to denote the ith sub-interval $[x_{i-1}, x_i]$ as its sub-length $x_i - x_{i-1}$.

Thus,

$$\Delta x_i = x_i - x_{i-1} \ (i = 1, 2, ..., n)$$

Let f be a *bounded real-valued* function on $[a, b]$. Evidently f is bounded on each sub-interval corresponding to each partition P. Let M_i, m_i be the bounds (supremum and infimum) of f in Δx_i.

From the two sums,

$$U(P, f) = \sum_{i=1}^{n} M_i \Delta x_i = M_1 \Delta x_1 + M_2 \Delta x_2 + \ldots + M_n \Delta x_n$$

$$L(P, f) = \sum_{t=1}^{n} M_i \Delta x_i = m_1 \Delta x_1 + b = m_2 \Delta x_2 + \ldots + m_n \Delta x_n$$

respectively called the *Upper* and the *Lower* (Darboux) *sums* of f corresponding to the partition P.

If M, m are the bounds of f in $[a, b]$, we have

$$m \le m_i \le M_i \le M$$

$$\Rightarrow \qquad m\,\Delta x_i \le m_i \Delta x_i \le M_i \Delta x_i \le M\,\Delta x_i$$

Putting $i = 1, 2, \ldots, n$ and adding all the inequalities, we get

$$m(b-a) \le L(P, f) \le U(P, f) \le M(b-a),\ b \ge a \qquad \ldots(1)$$

Now each partition gives rise to a pair of sums, the upper and the lower sums. By considering all partitions of $[a, b]$, we get a set U of upper sums and a set L of lower sums. The inequality (1) shows that both these sets are bounded and so each set has the supremum and the infimum. The *infimum* of the set of upper sums is called the *Upper integral* and the supremum of the set of lower sums is called the *Lower integral* over $[a, b]$. Thus

$$\overline{\int_a^b} f\,dx = \inf U \quad \text{or} \quad \inf\{U(P, f) : P \text{ is a partition of } [a, b]\}$$

$$\overline{\int_a^b} f\,dx = \sup \mathrm{L} \quad \text{or} \quad \sup\{L(P, f) : P \text{ is a partition of } [a, b]\}$$

These two integrals may or may not be equal.

Definition 1 (*Darboux's condition of integrability*). When the two integrals are equal, *i.e.*,

$$\overline{\int_a^b} f dx = \underline{\int_a^b} f dx = \int_a^b f dx$$

we say that f is *Riemann Integrable* (or simply *integrable*) over $[a, b]$ and the common value of these integrals is called the *Riemann Integral* (or simply the *integral*) of f over $[a, b]$.

The fact that f is integrable over $[a, b]$, we express by writing

$$f \in R[a, b] \text{ or } R \text{ simply.}$$

Evidently from equation (1),

$$m(b-a) \le \int_a^b f\,dx \le M(b-a),\ b \ge a \qquad \ldots(2)$$

Thus, the upper and lower integrals are defined for *every bounded function* but they may not necessarily be equal for every bounded function. There exist functions for which these integrals are not equal, such functions are not integrable. Thus the question of their equality, and hence the question of the integrability of a function, is a more delicate one and will be our main concern in the next few pages.

Note: For the sake of convenience, whenever the scope for confusion is not there, we shall omit the limits of integration and write simply,

$$\overline{\int} f\,dx,\quad \underline{\int} f\,dx,\quad \int f\,dx$$

Remarks:

1. The statement that $\int_a^b f\,dx$ exists, implies that the function f is *bounded* and *integrable* over $[a, b]$.
2. We have introduced the concept of integrability of a function subject to two very important limitations, *viz*. (*i*) the function is bounded, (*ii*) the interval is finite.
3. From equations (1) and (2), it follows that when $b > a$,

$$m(b-a) \le L(P,f) \le \int_a^b f\,dx \le U(P,f) \le M(b-a)$$

4. Since the upper integral is the greatest lower bound of the set of upper sums, therefore corresponding to any $\varepsilon_1 > 0$, $\exists$ an upper sum (or $\exists$ a partition P_1) such that

$$U(P_1 f) < \overline{\int_a^b} f\,dx + \varepsilon_1$$

Similarly,

$$L(P_2 f) > \underline{\int_a^b} f\,dx - \varepsilon_2$$

5. $U(P,f) - L(P,f) = \sum_i M_i\,\Delta x_i - \sum_i m_i\,\Delta x_i = \sum_i (M_i - m_i)\,\Delta x_i$

$(M_i - m_i)$ being the oscillation of f in the sub-interval Δx_i, $U(n, f) - L(P,f)$ is called the *oscillatory sum* denoted by $\omega(P,f)$ and is non-negative.

Example 1. Show that a constant function k is integrable and

$$\int_a^b k\,dx = k(b-a)$$

■ For any partition P of the interval $[a, b]$, we have

$$\begin{aligned} L(P,f) &= k\,\Delta x_1 + k\,\Delta x_2 + \dots + k\,\Delta x_n \\ &= k(\Delta x_1 + \Delta x_2 + \dots + \Delta x_n) \\ &= k(b-a) \end{aligned}$$

$$\Rightarrow \qquad \underline{\int_a^b} k\,dx = \sup L(P,f) = k(b-a)$$

$$\begin{aligned} \overline{\int_a^b} k\,dx &= \inf U(P,f) \\ &= \inf (k\,\Delta x_1 + k\,\Delta x_2 + \dots + k\,\Delta x_n) \\ &= k(b-a) \end{aligned}$$

Thus,

$$\underline{\int_a^b} k\,dx = \overline{\int_a^b} k\,dx = k(b-a)$$

which implies that the function k is integrable and

$$\int_a^b k\,dx = k(b-a)$$

Example 2. Show that the function f defined by

$$f(x) = \begin{cases} 0, & \text{when } x \text{ is rational,} \\ 1, & \text{when } x \text{ is irrational} \end{cases}$$

is not integrable on any interval.

■ Let us consider a partition P of an interval $[a, b]$.

$$U(P,f) = \sum_{i=1}^{n} M_1 \Delta x_1$$

$$= 1\,\Delta x_1 + 1\,\Delta x_2 + \ldots + 1\,\Delta x_n = b - a$$

$$\overline{\int_a^b} f\,dx = \inf U(P,f) = b - a$$

$$\underline{\int_a^b} f\,dx = \sup L(P,f)$$

$$= \sup\{0\Delta x_1 + 0\Delta x_2 + \ldots + 0\Delta x_n\} = 0$$

Thus,

$$\overline{\int} f\,dx \neq \underline{\int} f\,dx$$

Hence, the function f is not integrable.

Example 3. Show that x^2 is integrable on any interval $[0, k]$.

■ Let us consider the partition P of $[0, k]$ obtained by dividing the interval into n equal parts. Thus $[0, k/n, 2k/n, \ldots, nk/n]$ is the partition P, $[(i-1)(k/n)]^2$ and $[i(k/n)]^2$ are the lower and upper bounds of the function in Δx_i, and the length of each such interval is k/n.

$$\therefore \qquad U(P,x^2) = \frac{k^3}{n^3}(1^2 + 2^2 + 3^2 + \ldots + n^2)$$

$$= \frac{k^3}{n^3} \cdot \frac{n}{6}(n+1)(2n+1)$$

$$= \frac{k^3}{6}\left(1 + \frac{1}{n}\right)\left(2 + \frac{1}{n}\right)$$

and

$$L(P, x^2) = \frac{k^3}{n^3}\{0 + 1^2 + 2^2 + \ldots + (n-1)^2\}$$
$$= \frac{k^3}{6}\left(1 - \frac{1}{n}\right)\left(2 - \frac{1}{n}\right)$$

$\therefore$
$$\inf U(P, x^2) = \frac{k^3}{3} = \sup L(P, x^2)$$

Hence, the function x^2 is integrable and

$$\int_0^k x^2\, dx = k^3/3$$

Ex. 1. Show that $(3x + 1)$ is integrable on $[1, 2]$ and

$$\int_1^2 (3x+1)\, dx = \frac{11}{2}$$

Ex. 2. A function f is bounded on $[a, b]$. Show that

(*i*) when k is a positive constant,

$$\overline{\int_a^b} kf\, dx = k\overline{\int_a^b} f\, dx;\ \underline{\int_a^b} kf\, dx = k\underline{\int_a^b} f\, dx$$

and (*ii*) when k is a negative constant,

$$\overline{\int_a^b} kf\, dx = k\underline{\int_a^b} f\, dx;\ \underline{\int_a^b} kf\, dx = k\overline{\int_a^b} f\, dx$$

Deduce that if f is bounded and integrable over $[a, b]$, then so is kf, where k is any constant and then

$$\int_a^b kf\, dx = k\int_a^b f\, dx$$

[*Hint*: If M_i, m_i M_i, m_i are bounds of f in Δx_1, then $k M_i$, km_i, are bound of kf in Δx_i, if k is positive. But if k is negative, the bounds of kf are km_i, kM_i]

1.1 A Definition

The meaning of $\int_a^b f\, dx$ when $b \leq a$. If f is bounded and integrable on $[b, a]$, for $a > b$, we define

$$\int_a^b f\, dx = -\int_b^a f\, dx, \text{ when } a > b$$

Also

$$\int_a^b f\,dx = 0 \quad \text{when } a = b$$

1.2 Inequalities for Integrals

In an earlier section we have proved that for a bounded integrable function f,

$$m(b-a) \le \int_a^b f\,dx \le M(b-a), \text{ when } b \ge a \qquad \ldots(3)$$

If $b < a$, so that $a > b$, then as proved above

$$m(a-b) \le \int_b^a f\,dx \le M(a-b), \text{ when } a > b$$

$$\Rightarrow \qquad -m(a-b) \ge -\int_b^a f\,dx \ge -M(a-b)$$

$$\Rightarrow \qquad m(b-a) \ge \int_a^b f\,dx \ge M(b-a), \text{ when } b < a \qquad \ldots(4)$$

We shall now make some interesting deductions from these two inequalities.

Deduction 1. If f is bounded and integrable on $[a, b]$, then there exists a number λ lying between the bounds of f such that

$$\int_a^b f\,dx = \lambda(b-a)$$

Deduction 2. If f is continuous and integrable on $[a, b]$, then there exists a number c between a and b such that

$$\int_a^b f\,dx = (b-a)\,f(c)$$

Deduction 3. If f is bounded and integrable on $[a, b]$ and k is a number such that $|f(x)| \le k$ for all $x \in [a, b]$, then

$$\left| \int_a^b f\,dx \right| \le k\,|b-a|$$

Let M, m be the bounds of $f(x)$. Now

$$|f(x)| \le k, \;\; \forall\; x \in [a, b]$$

$$\therefore \qquad -k \le f(x) \le k \;\Rightarrow\; -k \le m \le f(x) \le M \le k$$

which for $b \ge a$, implies that

$$-k(b-a) \le m(b-a) \le \int_a^b f\,dx \le M(b-a) \le k(b-a)$$

$$\Rightarrow \qquad \left|\int_a^b f\,dx\right| \le k(b-a)$$

If $b < a$, so that $a > b$, we have

$$\left|\int_b^a f\,dx\right| \le k\,|a-b| \quad \Rightarrow \quad \left|\int_a^b f\,dx\right| \le k\,|b-a|$$

The result is trivial for $a = b$.

Deduction 4. If f is bounded and integrable on $[a, b]$, and $f(x) \ge 0$ for all $x \in [a, b]$, then

$$\int_a^b f\,dx \begin{cases} \ge 0, & \text{if } b \ge a \\ \le 0, & \text{if } b \le a \end{cases}$$

Since $f(x) \ge 0$, for all $x \in [a, b]$, therefore, the lower bound $m \ge 0$.

The result follows from the inequalities (3) and (4) above.

Deduction 5. If f, g are bounded and integrable on $[a, b]$, such that $f \ge g$, then

$$\int_a^b f\,dx \ge \int_a^b g\,dx \quad \text{when } b \ge a$$

and

$$\int_a^b f\,dx \le \int_a^b g\,dx \quad \text{when } b \le a$$

Now

$$f \ge g \Rightarrow f - g \ge 0, \quad \forall\, x \in [a, b]$$

Hence using Deduction 4, we have

$$\int_a^b (f-g)\,dx \ge 0 \qquad \text{if} \quad b \ge a$$

or

$$\int_a^b f\,dx \ge \int_a^b g\,dx \quad \text{if} \quad b \ge a$$

Similarly

$$\int_a^b f\,dx \le \int_a^b g\,dx \quad \text{if} \quad b \ge a$$

Note: We have assumed the result $\int (f - g)\,dx = \int f\,dx - \int g\,dx$. It will however be proved in § 5, Theorem 6.

2. REFINEMENT OF PARTITIONS

Definition. For any partition P, the length of the largest sub-interval is called the *norm* or *mesh* of the partition and is denoted as $\mu(P)$ (or simply μ).

$$\therefore \qquad \mu(P) = \max \Delta x_i \ (1 \le i \le n)$$

A partition P^* is said to be a *refinement* of P if $P^* \supseteq P$, *i.e.*, every point of P is a point of P^*.

We also say that P^* *refines* P or that P^* is *finer than* P.

If P_1 and P_2 are two partitions, then we say that P^* is their *common refinement* if $P^* = P_1 \cup P_2$.

Theorem 1. *If P^* is a refinement of a partition P, then for a bounded function f,*

(*i*) $L(P^*, f) \ge L(P, f)$, *and*

(*ii*) $U(P^*, f) \le U(P, f)$.

To prove (*i*), suppose first that P^* contains just one point more than P.

Let this extra point be ξ, and suppose that this point is in Δx_i, that is, $x_{i-1} < \xi < x_i$.

As the function is bounded over the entire interval $[a, b]$, it is bounded in every sub-interval Δx_i $(i = 1, 2, \ldots, n)$. Let w_1, w_2, m_i be the infimum (g.l.b.) of f in the intervals $[x_{i-1}, \xi], [\xi, x_i], [x_{i-1}, x_i]$, respectively.

Clearly $m_i \le w_1, m_i \le w_2$.

$$\begin{aligned} \therefore \quad L(P^*, f) - L(P, f) &= w_1(\xi - x_{i-1}) + w_2(x_i - \xi) - m_i(x_i - x_{i-1}) \\ &= (w_1 - m_i)(\xi - x_{i-1}) + (w_2 - m_i)(x_i - \xi) \ge 0 \end{aligned}$$

($\because$ each bracket is positive)

If P^* contains p points more than P, we repeat the above reasoning p times and arrive at

$$L(P^*, f) \ge L(P, f)$$

Similarly, we can prove that

$$U(P^*, f) \le U(P, f)$$

Corollary. *If a refinement P^* of P contains p points more than P, and $|f(x)| \le k$, for all $x \in [a, b]$, then*

$$L(P, f) \le L(P^*, f) \le L(P, f) + 2pk\mu$$
$$U(P, f) \ge U(P^*, f) \ge U(P, f) - 2pk\mu$$

Proceeding as in the above theorem, when P^* contains one point more than P, we get

$$L(P^*, f) - L(P, f) = (w_1 - m_i)(\xi - x_{i-1}) + (w_2 - m_i)(x_i - \xi)$$

Since $|f(x)| \le k$, for all $x \in [a, b]$, therefore

$$-k \le m_i \le w_1 \le k$$

$$\Rightarrow \qquad 0 \le w_1 - m_i \le 2k$$

Similarly,

$$0 \le w_2 - m_i \le 2k$$

$$\therefore \quad L(P^*, f) - L(P, f) \le 2k(\xi - x_{i-1}) + 2k(x_i - \xi)$$

$$= 2k\ \Delta x_i$$

$$\le 2k\mu, \text{ where } \mu \text{ is norm of } P$$

Now supposing that each additional point is introduced one by one, by repeating the above reasoning p times, we get

$$L(P^*, f) \le L(P, f) + 2pk\mu$$

Also
$$L(P, f) \le L(P^*, f)$$

$$\therefore \quad L(P, f) \le L(P^*, f) \le L(P, f) + 2pk\mu$$

Similarly, the other result may be proved.

Ex. If P^* is a refinement of P, then

$$U(P^*, f) - L(P^*, f) \le U(P, f) - L(P, f)$$

Theorem 2. *For any two partitions P_1, P_2,*

$$L(P_1, f) \le U(P_2, f)$$

i.e., no upper sum can ever be less than any lower sum.

Let P^* be a common refinement of P_1, P_2, so that

$$P^* = P_1 \cup P_2$$

Using Theorem 1, we get

$$L(P_1, f) \le L(P^*, f) \le U(P^*, f) \le U(P_2, f) \qquad ...(1)$$

$$\therefore \quad L(P_1, f) \le U(P_2, f)$$

Corollary. *For any bounded function f,*

$$\underline{\int} f\, dx \le \overline{\int} f dx$$

By keeping P_2 fixed and taking the l.u.b. over all partitions P_1, (1) of Theorem 2 gives

$$\underline{\int} f\, dx \le U(P_2, f) \qquad ...(2)$$

Taking the g.l.b. over all P_2 in equation (2), we get

$$\underline{\int} f\, dx \le \overline{\int} f\, dx$$

3. DARBOUX'S THEOREM

If f is a bounded function on $[a, b]$, *then to every* $\varepsilon > 0$, *there corresponds* $\delta > 0$ *such that*

(*i*) $U(P, f) < \overline{\int_a^b} f\, dx + \varepsilon$,

(*ii*) $L(P, f) > \underline{\int_a^b} f\, dx - \varepsilon$

for every partition P of $[a, b]$ *with norm* $\mu(P) < \delta$.

Let us first attend to (*i*).

As f is bounded, therefore $\exists\, k > 0$, such that

$$|f(x)| \le k, \ \forall\, x \in [a, b]$$

Again, since the upper integral is the infimum (g.l.b.) of the set of upper sums, therefore to every $\varepsilon > 0$ there exists a partition $P_1 = \{x_0, x_1, x_2, ..., x_p\}$ of $[a, b]$ such that

$$U(P_1, f) < \overline{\int_a^b} f dx + \frac{1}{2}\varepsilon \qquad ...(1)$$

The partition P_1 has $p - 1$ points besides $x_0 (= a)$ and $x_p (= b)$.

Let δ be a positive number such that $2k(p-1)\delta = \frac{1}{2}\varepsilon$ ...(2)

Let P be any partition with norm $\mu(P) < \delta$.

Let, further, P^* be a refinement of P and P_1, so that $P^* = P \cup P_1$.

As P^* is a refinement of P having at the most $p - 1$ more points than P, therefore (using Corollary, Theorem 1), we get

$$\begin{aligned} U(P, f) - 2k(p-1)\delta &\le U(P^*, f) \\ &\le U(P_1, f) \\ &< \overline{\int_a^b} f dx + \frac{1}{2}\varepsilon \qquad \text{[using equation (1)]} \end{aligned}$$

$$\Rightarrow \qquad U(P, f) < \overline{\int_a^b} f\, dx + \frac{1}{2}\varepsilon - \frac{1}{2}\varepsilon = \overline{\int_a^b} f\, dx + \varepsilon \qquad \text{[using equation (2)]}$$

for any partition P with norm $\mu(P) < \delta$.

Similarly, we can prove the other part.

Note: The definition of infimum also leads to the fact that

$$U(P, f) < \overline{\int_a^b} f\, dx + \varepsilon$$

but this implies that for every $\varepsilon > 0$ there exists at least one partition P with this property. The importance of Darboux's Theorem lies in the fact that it asserts the existence of an infinite number of partitions which have this property, with the only restriction that their norm $\mu(P) < \delta$, depends on the choice of ε.

4. CONDITIONS OF INTEGRABILITY

We have stated earlier that a bounded function is said to be integrable when the upper and the lower integrals are equal. We now formalise and give the necessary and sufficient conditions for the integrability of a function in two forms.

Theorem 3. ***First form.*** *A necessary and sufficient condition for the integrability of a bounded function f is that to every* $\varepsilon > 0$*, there corresponds* $\delta > 0$ *such that for every partition P of* $[a, b]$ *with norm* $\mu(P) < \delta$,

$$U(P, f) - L(P, f) < \varepsilon$$

The condition is necessary. The bounded function f is integrable,

$$\therefore \qquad \underline{\int_a^b} f\, dx = \overline{\int_a^b} f\, dx = \int_a^b f\, dx$$

Let ε be any positive number. By Darboux's Theorem there exists $\delta > 0$ such that for every partition P with norm $\mu(P) < \delta$,

$$U(P,f) < \overline{\int_a^b} f\, dx + \frac{1}{2}\varepsilon = \int_a^b f\, dx + \frac{1}{2}\varepsilon \qquad \text{...(1)}$$

$$L(P,f) > \underline{\int_a^b} f\, dx - \frac{1}{2}\varepsilon = \int_a^b f\, dx - \frac{1}{2}\varepsilon \qquad \text{...(2)}$$

or

$$-L(P,f) < -\int_a^b f\, dx + \frac{1}{2}\varepsilon \qquad \text{...(3)}$$

From equations (1) and (3), we get on adding

$$U(P, f) - L(P, f) < \varepsilon$$

for every partition P with norm $\mu(P) < \delta$.

The condition is sufficient. Let ε be any positive number. For any partition P with norm $\mu(P) < \delta$ (depending on ε), we are given that

$$U(P, f) - L(P, f) < \varepsilon$$

Also for any partition P, we know that

$$L(P,f) \le \underline{\int_a^b} f\, dx \le \overline{\int_a^b} f\, dx \le U(P, f)$$

$$\Rightarrow \qquad \overline{\int_a^b} f\, dx - \underline{\int_a^b} f\, dx \le U(P,f) - L(P,f) < \varepsilon$$

Since ε is an arbitrary positive number, therefore we see that a non-negative number is less than every positive number. Hence it must be equal to zero,

$$\Rightarrow \qquad \overline{\int_a^b} f\, dx - \underline{\int_a^b} f\, dx = 0$$

$$\Rightarrow \qquad \overline{\int_a^b} f\, dx = \underline{\int_a^b} f\, dx$$

so that f is integrable.

Note: The theorem can also be stated as follows:
A necessary and sufficient condition for the integrability of a bounded function f is that

$$\lim \{U(P,f) - L(P,f)\} = 0$$

when the norm $\mu(P)$ of the partition P tends to 0.

Theorem 4. ***Second form.*** *A bounded function f is integrable on $[a, b]$ iff for every $\varepsilon > 0$ there exists a partition P of $[a, b]$, such that*

$$U(P,f) - L(P,f) < \varepsilon$$

The condition is necessary. Suppose the function f is integrable, so that

$$\underline{\int_a^b} f\, dx = \overline{\int_a^b} f\, dx = \int_a^b f\, dx$$

Let ε be a positive number.

Since the upper and the lower integrals are the infimum and the supremum, respectively, of the upper and the lower sums, therefore $\exists$ partitions P_1 and P_2 such that

$$U(P_1,f) < \overline{\int_a^b} f\, dx + \frac{1}{2}\varepsilon = \int_a^b f\, dx + \frac{1}{2}\varepsilon$$

$$L(P_2,f) > \underline{\int_a^b} f\, dx - \frac{1}{2}\varepsilon = \int_a^b f\, dx - \frac{1}{2}\varepsilon$$

Let P be the common refinement of P_1 and P_2, *i.e.*, $P = P_1 \cup P_2$,

$$\therefore \qquad U(P,f) \le U(P_1,f) < \int_a^b f\, dx + \frac{1}{2}\varepsilon < L(P_2,f) + \varepsilon$$

$$\le L(P,f) + \varepsilon$$

Thus $\exists$ a partition P such that

$$U(P,f) - L(P,f) < \varepsilon$$

The condition is sufficient. Let ε be any positive number.

Let P be a partition for which

$$U(P,f) - L(P,f) < \varepsilon$$

For any partition P, we know that

$$L(P,f) \le \underline{\int_a^b} f\, dx \le \overline{\int_a^b} f\, dx \le U(P,f)$$

$$\therefore \qquad \overline{\int_a^b} f\, dx - \underline{\int_a^b} f\, dx \le U(P,f) - L(P,f) < \varepsilon$$

The non-negative number, being less than every positive number ε, must be zero.

$$\therefore \qquad \overline{\int_a^b} f\, dx = -\underline{\int_a^b} f\, dx$$

so that f is integrable.

Note: Comparison of the two forms indicates that from the necessary point of view, the first form is stronger than the second but from the sufficiency view point the second form is stronger than the first.

4.1 Deduction 6. *A function f is integrable over $[a, b]$ iff there is a number I lying between $L(P, f)$ and $U(P, f)$ such that for any $\varepsilon > 0$, $\exists$ a partition P of $[a, b]$ such that*

$$|U(P, f) - I| < \varepsilon, \text{ and } |I - L(P, f)| < \varepsilon.$$

Necessary: As $f \in R[a, b]$ therefore for $\varepsilon > 0$, $\exists$ a partition P of $[a, b]$ such that

$$|U(P, f) - L(P, f)| < \varepsilon$$

If I is a number between $L(P, f)$ and $U(P, f)$, then

$$|U(P, f) - I| < |U(P, f) - L(P, f)| < \varepsilon$$

and

$$|I - L(P, f)| < |U(P, f) - L(P, f)| < \varepsilon$$

Hence, the result.

Sufficient: For $\varepsilon > 0$, $\exists$ a partition P such that

$$|U(P, f) - I| < \frac{1}{2}\varepsilon \text{ and } |I - L(P, f)| < \frac{1}{2}\varepsilon$$

$$\therefore \quad |U(P, f) - L(P, f)| \leq |U(P, f) - I| + |I - L(P, f)|$$

$$< \frac{1}{2}\varepsilon + \frac{1}{2}\varepsilon = \varepsilon$$

$$\Rightarrow \quad f \in R[a, b]$$

Deduction 7. A function f is integrable over $[a, b]$ iff there is a number I such that for any $\varepsilon > 0, \exists\, \delta$ such that for all partitions P with mesh $\mu(P) < \delta$,

$$|U(P, f) - I| < \varepsilon, \text{ and } |I - L(P, f)| < \varepsilon$$

The proof is similar to that of deduction 6 and is left to the reader.

Note: We know that $U(P^*, f) < U(P, f)$ when $P^* \supset P$. Therefore, the upper sum becomes smaller and smaller as the partition gets finer and finer. Thus the upper integral $\overline{\int_a^b} f\, dx$, which is the g.l.b. of the set U of upper sums, can be looked upon as $\lim_{\mu(P)\to 0} U(P, f)$ or that for $\varepsilon > 0, \exists\, \delta > 0$ such that for all partitions P with mesh $\mu(P) < \delta, |U(P, f) - \overline{\int} f\, dx| < \varepsilon$. Thus Deduction 7 merely states that the upper and the lower sums converge to the same quantity I and moreover $I = \int_a^b f\, dx$.

CHAPTER 9

5. INTEGRABILITY OF THE SUM AND DIFFERENCE OF INTEGRABLE FUNCTIONS

Theorem 5. *If f_1 and f_2 are two bounded and integrable functions on $[a, b]$, then $f = f_1 + f_2$ is also integrable on $[a, b]$, and*

$$\int_a^b f_1\, dx + \int_a^b f_2\, dx = \int_a^b f\, dx$$

Clearly f is bounded on $[a, b]$.

Let $P = \{a = x_0, x_1, \ldots, x_n = b\}$ be any partition of $[a, b]$ and $M_i', m_i'; M_i'', m_i'', M_i, m_i$ be the bounds of f_1, f_2 and f respectively in Δx_i.

[$M_i' + M_i''$ and $m_i' + m_i''$ are rough upper and rough lower bounds while M_i and m_i, the supremum and the infimum of f in Δx_i.]

$$m_i' + m_i'' \le m_i \le M_i \le M_i' + M_i'' \qquad \ldots(1)$$

Multiplying by Δx_i and adding all these inequalities for $i = 1,2,3,\ldots n$, we get

$$L(P, f_1) + L(P, f_2) \le L(P, f) \le U(P, f) \le U(P, f_1) + U(P, f_2) \qquad \ldots(2)$$

Let ε be any positive number.

Since f_1, f_2 are integrable therefore, we can choose $\delta > 0$ such that, for any partition P with norm $\mu(P) < \delta$, we have

$$\left.\begin{aligned} U(P, f_1) - L(P, f_1) < \frac{1}{2}\varepsilon \\ U(P, f_2) - L(P, f_2) < \frac{1}{2}\varepsilon \end{aligned}\right\} \qquad \ldots(3)$$

Thus for any partition P with norm $\mu(P) < \delta$ we have, from (2) and (3),

$$\begin{aligned} U(P, f) - L(P, f) &\le U(P, f_1) + U(P, f_2) - L(P, f_1) - L(P, f_2) \\ &< \frac{1}{2}\varepsilon + \frac{1}{2}\varepsilon = \varepsilon \end{aligned}$$

Thus, the function f is integrable.

Now we proceed to prove the second part.

Since f_1, f_2 are integrable and ε is any positive number, therefore, by Darboux's Theorem, $\exists\, \delta > 0$ such that for all partitions P whose norm $\mu(P) < \delta$, we have

$$U(P, f_1) < \int_a^b f_1\, dx + \frac{1}{2}\varepsilon, \text{ and } U(P, f_2) < \int_a^b f_2 dx + \frac{1}{2}\varepsilon \qquad \ldots(4)$$

Also,

$$\int_a^b f\, dx \le U(P, f) \le U(P, f_1) + U(P, f_2) \qquad \text{[using (2)]}$$

$$< \int_a^b f_1\,dx + \int_a^b f_2\,dx + \varepsilon \qquad \text{[using (4)]}$$

Since ε is arbitrary, we conclude that

$$\int_a^b f\,dx \le \int_a^b f_1\,dx + \int_a^b f_2\,dx \qquad ...(5)$$

Proceeding with $(-f_1),(-f_2)$ in place of f_1, f_2 respectively, we get

$$\int_a^b (-f)\,dx \le \int_a^b (-f_1)\,dx + \int_a^b (-f_2)\,dx$$

or

$$\int_a^b f\,dx \ge \int_a^b f_1\,dx + \int_a^b f_2\,dx \qquad ...(6)$$

From equations (5) and (6),

$$\int_a^b f\,dx = \int_a^b f_1\,dx + \int_a^b f_2\,dx$$

Note: When f_1 is integrable, for $\varepsilon > 0, \exists\, \delta_1 > 0$ such that for $\mu(P) < \delta_1, U(P, f) - L(P, f_1) < \frac{1}{2}\varepsilon$. Similarly $\exists\, \delta_2 > 0$ for the functions f_2. However, $\delta = \min(\delta_1, \delta_2)$ works for both the functions. It is this δ which was used for (3) above, δ for (4) was also selected by a similar reasoning.

Ex. Prove the above theorem by using Theorem 4 for condition of integrability.

Theorem 6. *If f_1, f_2 are two bounded and integrable functions on $[a, b]$, then $f = f_1 - f_2$ is also integrable on $[a, b]$ and*

$$\int_a^b f\,dx = \int_a^b f_1\,dx - \int_a^b f_2\,dx$$

Let $f = f_1 + (-f_2)$, so that f is bounded on $[a, b]$.

Let $P = \{a = x_0, x_1, x_2, \ldots, x_n = b\}$ be any partition of $[a, b]$ and $M_i', m_i'; M_i'', m_i''; M_i, m_i$ be the bounds of f_1, f_2 and f respectively in Δx_i. Clearly the bounds of $(-f_2)$ are $-m_i''$ and $-M_i''$.

$$\therefore \qquad m_i' - M_i'' \le m_i \le M_i \le M_i' - m_i'' \qquad ...(1)$$

Multiplying by Δx_i and adding all these inequalities for $i = 1, 2, ..., n$, we get

$$L(P, f_1) - U(P, f_2) \le L(P, f) \le U(P, f) \le U(P, f_1) - L(P, f_2) \qquad ...(2)$$

Let $\varepsilon > 0$ be a given number.

Since f_1, f_2 are integrable, therefore $\exists\, \delta > 0$ such that for any partition P whose norm $\mu(P) < \delta$, we have

$$\left.\begin{aligned} U(P, f_1) - L(P, f_1) < \frac{1}{2}\varepsilon \\ U(P, f_2) - L(P, f_2) < \frac{1}{2}\varepsilon \end{aligned}\right\} \qquad ...(3)$$

Thus for any partition P with norm $\mu(P) < \delta$, we have from equations (2) and (3),

$$U(P, f) - L(P, f) \le U(P, f_1) - L(P, f_2) - L(P, f_1) + U(P, f_2)$$

$$< \frac{1}{2}\varepsilon + \frac{1}{2}\varepsilon = \varepsilon$$

Thus, the function f is integrable.

Let us now prove the second part.

Since f_1, f_2 are integrable and ε is any positive number, therefore, by Darboux's Theorem $\exists\ \delta > 0$ such that for all partitions P with norm $\mu(P) < \delta$, we have

$$U(P, f_1) < \int_a^b f_1 dx + \frac{1}{2}\varepsilon, \text{ and } L(P, f_2) > \int_a^b f_2\, dx - \frac{1}{2}\varepsilon \qquad ...(4)$$

$$\therefore \qquad \int_a^b f\, dx \le U(P, f)$$

$$\le U(P, f_1) - L(P, f_2) \qquad \text{[using (8)]}$$

$$< \int_a^b f_1\, dx - \int_a^b f_2\, dx + \varepsilon$$

Since ε is arbitrary, we conclude that

$$\int_a^b f\, dx \le \int_a^b f_1\, dx - \int_a^b f_2\, dx$$

Proceeding with $(-f_1)$ and $(-f_2)$ in place of f_1 and f_2 respectively, we get

$$\int_a^b f\, dx \ge \int_a^b f_1\, dx - \int_a^b f_2\, dx$$

$$\therefore \qquad \int_a^b f\, dx = \int_a^b f_1\, dx - \int_a^b f_2\, dx$$

Ex. Prove the above theorem by using Theorem 4.

Theorem 7. (*i*) *If a bounded function f is integrable on* $[a, b]$, *then it is also integrable on* $[a, c]$ *and* $[c, b]$, *where c is a point of* $[a, b]$.

(*ii*) *Conversely, if f is bounded and integrable on* $[a, c]$, $[c, b]$, *then it is also integrable on* $[a, b]$.

(*iii*) *Also in either case,*

$$\int_a^b f\, dx = \int_a^c f\, dx + \int_c^b f\, dx, \quad a \le c \le b$$

(*i*) Since $f \in R[a, b]$ therefore for $\varepsilon > 0, \exists$ a partition P such that

$$U(P,f) - L(P,f) < \varepsilon$$

Let P^* be a refinement of P such that

$$P^* = P \cup \{c\}$$

$\therefore$ $$L(P,f) \le L(P^*,f) \le U(P^*,f) \le U(P,f) \quad ...(1)$$

$\Rightarrow$ $$U(P^*,f) - L(P^*,f) \le U(P,f) - L(P,f)$$

$$< \varepsilon \quad ...(2)$$

Let P_1, P_2 denote the sets of points of P^* between $[a, c]$, $[c, b]$ respectively. Clearly P_1, P_2 are partitions of $[a, c]$, $[c, b]$ respectively and $P^* = P_1 \cup P_2$.

Also

$$U(P^*,f) = U(P_1,f) + U(P_2,f) \quad ...(3)$$

and

$$L(P^*,f) = L(P_1,f) + L(P_2,f) \quad ...(4)$$

$\therefore$ $$\{U(P_1,f) - L(P_1,f)\} + \{U(P_2,f) - L(P_2,f)\} = U(P^*,f) - L(P^*,f)$$

$$< \varepsilon$$

Since each bracket on the left is non-negative, it follows that partitions P_1, P_2 exist such that

$$U(P_1,f) - L(P_1,f) < \varepsilon/2$$
$$U(P_2,f) - L(P_2,f) < \varepsilon/2$$

$\Rightarrow$ Integrable f is on $[a, c]$ and $[c, b]$

(*ii*) Let $f \in R$ over $[a, c]$ and $[c, b]$.

Therefore for $\varepsilon > 0$, we can find partitions P_1, P_2 of $[a, c]$, $[c, b]$ respectively such that

$$U(P_1,f) - L(P_1,f) < \frac{1}{2}\varepsilon \quad ...(5)$$

$$U(P_2,f) - L(P_2,f) < \frac{1}{2}\varepsilon \quad ...(6)$$

Let $P^* = P_1 \cup P_2$.

Clearly P^* is a partition of $[a, b]$.

$\therefore$ $$U(P^*,f) - L(P^*,f) = U(P_1,f) + U(P_2,f) - L(P_1,f) - L(P_2,f)$$

$$< \frac{1}{2}\varepsilon + \frac{1}{2}\varepsilon = \varepsilon$$

Thus for $\varepsilon > 0, \exists$ a partition P^* of $[a, b]$ such that

$$U(P^*,f) - L(P^*,f) < \varepsilon.$$

$\Rightarrow$ $$f \in R \text{ over } [a, b]$$

(*iii*) We know that for any functions S and T, if $W = S + T$, then $\inf W > \inf S + \inf T$

Now for any partitions P_1, P_2 of $[a, c]$, $[c, b]$ respectively, if $P^* = P_1 \cup P_2$, then

$$U(P^*,f) = U(P_1,f) + U(P_2,f)$$

CHAPTER 9

Hence, on taking the infimum for all partitions, we get

$$\int_a^b f\,dx \ge \overline{\int_a^c} f\,dx + \overline{\int_a^b} f\,dx$$

But since f is integrable on $[a, c]$, $[c, b]$, $[a, b]$,

$$\therefore \qquad \int_a^b f\,dx \ge \int_a^c f\,dx + \int_c^b f\,dx \qquad ...(7)$$

Proceeding with $(-f)$ in place of f, we get

$$\int_a^b f\,dx \le \int_a^c f\,dx + \int_c^b f\,dx \qquad ...(8)$$

Equations (7) and (8), imply that

$$\int_a^b f\,dx = \int_a^c f\,dx + \int_c^b f\,dx$$

Ex. Do the above theorem by using Theorem 3 for conditions of integrability.

5.1* Integrability of the Product, Quotient and the Modulus of Integrable Functions

Before taking up the main theorems, let us prove a simple but useful lemma.

Lemma. *The oscillation of a bounded function f on an interval $[a, b]$ is the supremum of the set $\{|f(x_1) - f(x_2)| : x_1, x_2 \in [a, b]\}$ of numbers.*

Let M, m be the bounds of f, on $[a, b]$. Now

$$m \le f(x_1), f(x_2) \le M, \ \forall\, x_1, x_2 \in [a, b]$$

$$\Rightarrow \qquad |f(x_1) - f(x_2)| \le M - m, \ \forall\, x_1, x_2 \in [a, b] \qquad ...(1)$$

$\Rightarrow$ $M - m$ is an upper bound of the set in question.

Let $\varepsilon > 0$ be any given number.

Since M is supremum of f, therefore $\exists\, x' \in [a, b]$ such that

$$f(x') > M - \frac{1}{2}\varepsilon \qquad ...(2)$$

Similarly, $\exists\, x'' \in [a, b]$ such that

$$f(x'') < m + \frac{1}{2}\varepsilon \qquad ...(3)$$

Equations (2) and (3) imply that $\exists\, x', x'' \in [a, b]$ such that

$$f(x') - f(x'') > M - m - \varepsilon$$

$$\Rightarrow \qquad |f(x') - f(x'')| > M - m - \varepsilon \qquad ...(4)$$

(1) and (4) imply that $M - m$ is an upper bound and no number less than $M - m$ can be an upper bound of the set in question.

$$\therefore \qquad M - m = \sup \{|f(x_1) - f(x_2)| : x_1, x_2 \in [a, b]\}$$

Theorem 8*. *If f_1 and f_2 are two bounded and integrable functions on $[a, b]$, then their product $f_1 f_2$ is also bounded and integrable on $[a, b]$.*

Since f_1, f_2 are bounded therefore, $\exists\, k > 0$, such that for all $x \in [a, b]$,

$$|f_1(x)| \le k, \; |f_2(x)| \le k$$

$$\Rightarrow \qquad |(f_1 f_2)(x)| = |f_1(x) f_2(x)| \le k^2, \quad \forall\, x \in [a, b]$$

$$\Rightarrow \qquad f_1 \cdot f_2 \text{ is bounded}$$

Let $P = \{a = x_0, x_1, x_2, \ldots, x_n = b\}$ be any partition of $[a, b]$.

Let M_i', m_i'; M_i'', m_i'', M_i, m_i be the bounds of f_1, f_2 and $f_1 f_2$ respectively in Δx_i.

We have for all $x_1, x_2 \in \Delta x_i$,

$$\begin{aligned}(f_1 f_2)(x_2) - (f_1 f_2)(x_1) &= f_1(x_2) f_2(x_2) - f_1(x_1) f_2(x_1) \\ &= f_2(x_2)[f_1(x_2) - f_1(x_1)] + f_1(x_1)[f_2(x_2) - f_2(x_1)]\end{aligned}$$

$$\begin{aligned}\Rightarrow \quad |(f_1 f_2)(x_2) - (f_1 f_2)(x_1)| &\le |f_2(x_2)| \cdot |f_1(x_2) - f_1(x_1)| + |f_1(x_1)| \cdot |f_2(x_2) - f_2(x_1)| \\ &\le k(M_i' - m_i') + k(M_i'' - m_i'')\end{aligned}$$

$$\Rightarrow \qquad M_i - m_i \le k(M_i' - m_i') + k(M_i'' - m_i'') \qquad \ldots(1)$$

Now let $\varepsilon > 0$ be a given number.

Since f_1, f_2 are integrable, therefore $\exists\, \varepsilon > 0$ such that for any partition P with norm $\mu(P) < \delta$,

$$U(P, f_1) - L(P, f_1) < \varepsilon/2k, \quad U(P, f_2) - L(P, f_2) < \varepsilon/2k$$

Hence, from (1), multiplying by Δx_i and adding all such inequalities, we have for any partition P with norm $\mu(P) < \delta$,

$$\begin{aligned}U(P, f_1 f_2) - L(P, f_1 f_2) &\le k[U(P, f_1) - L(P, f_1)] + k[U(P, f_2) - L(P, f_2)] \\ &< k(\varepsilon/2k) + k(\varepsilon/2k) = \varepsilon\end{aligned}$$

implying that $f_1 f_2$ is integrable on $[a, b]$.

Theorem 9. *If f_1, f_2 are two bounded and integrable functions on $[a, b]$ and there exists a number $\lambda > 0$ such that $|f_2(x)| \ge \lambda$, for all x in $[a, b]$, then f_1 / f_2 is bounded and integrable on $[a, b]$.*

Since f_1, f_2 are bounded, therefore, $\exists$ positive number k such that

$$|f_1(x)| \le k, \; \lambda \le |f_2(x)| \le k, \quad \forall\, x \in [a, b]$$

$$\Rightarrow \qquad |(f_1/f_2)(x)| = |f_1(x)/f_2(x)| \le k/\lambda, \quad \forall\, x \in [a, b]$$

$$\Rightarrow \qquad f_1/f_2 \text{ is bounded}$$

Let $P = \{a = x_0, x_1, \ldots, x_n = b\}$ be a partition of $[a, b]$ and let M_i', m_i'; M_i'', m_i''; m_i, M_i be the bounds of $f_1, f_2, f_1/f_2$, respectively in Δx_i. We have for all $x_1, x_2 \in \Delta x_i$,

$$|(f_1/f_2)(x_2) - (f_1/f_2)(x_1)| = \left|\frac{f_1(x_2)}{f_2(x_2)} - \frac{f_1(x_1)}{f_2(x_1)}\right|$$

$$= \left| \frac{f_2(x_1)[f_1(x_2) - f_1(x_1)] - f_1(x_1)[f_2(x_2) - f_2(x_1)]}{f_2(x_2)\, f_2(x_1)} \right|$$

$$\leq (k/\lambda^2)|f_1(x_2) - f_1(x_1)| + (k/\lambda^2)|f_2(x_2) - f_2(x_1)|$$

$$\leq (k/\lambda^2)(M_i' - m_i') + (k/\lambda^2)(M_i'' - m_i'')$$

$$\Rightarrow \quad M_i - m_i \leq (k/\lambda^2)(M_i' - m_i') + (k/\lambda^2)(M_i'' - m_i'') \quad ...(1)$$

Now, let $\varepsilon > 0$ be a given number.

Since f_1, f_2 are integrable, therefore $\exists\ \delta > 0$ such that for any partition P with norm $\mu(P) < \delta$,

$$U(P, f_1) - L(P, f_1) < \varepsilon\lambda^2/2k$$

$$U(P, f_2) - L(P, f_2) < \varepsilon\lambda^2/2k$$

Hence from equation (1), for any partition P with $\mu(P) < \delta$, we have

$$U(P, f_1/f_2) - L(P, f_1/f_2) \leq (k/\lambda^2)\,[U(P, f_1) - L(P, f_1)]$$

$$+ (k/\lambda^2)\,[U(P, f_2) - L(P, f_2)]$$

$$< (k/\lambda^2)(\varepsilon\lambda^2/2k) + (k/\lambda^2)(\varepsilon\lambda^2/2k) = \varepsilon$$

implying that f_1/f_2 is integrable on $[a, b]$.

Theorem 10. *If f is bounded and integrable on $[a, b]$, then $|f|$ is also bounded and integrable on $[a, b]$. Moreover*

$$\left| \int_a^b f\, dx \right| \leq \int_a^b |f|\, dx$$

Since f is bounded, therefore $\exists\ k > 0$, such that

$$|f(x)| \leq k, \quad \forall\ x \in [a, b]$$

$\Rightarrow$ the function $|f|$ is bounded.

Again, since f is integrable, $\exists$ a partition $P = \{a = x_0, x_1, \ldots, x_n = b\}$ of $[a, b]$ such that

$$U(P, f) - L(P, f) < \varepsilon$$

Let $M_i, m_i;\ M_i', m_i'$ be the bounds of f and $|f|$ in Δx_i.

We have for all $x_1, x_2 \in \Delta x_i$

$$||f|(x_2) - |f|(x_1)| = ||f(x_2)| - |f(x_1)|| \leq |f(x_2) - f(x_1)| \leq M_i - m_i$$

$$\Rightarrow \quad M_i' - m_i' \leq M_i - m_i$$

This implies that for any partition P,

$$U(P, |f|) - L(P, |f|) \leq U(P, f) - L(P, f) < \varepsilon$$

Hence, $|f|$ is integrable on $[a, b]$.

Since $f(x), -f(x) \leq |f(x)| = |f|(x)$, for all x in $[a, b]$, therefore by Deduction 5, we have

$$\int_a^b f\, dx \leq \int_a^b |f|\, dx$$

and $$-\int_a^b f\,dx = \int_a^b (-f)\,dx \le \int_a^b |f|\,dx$$

$\Rightarrow$ $$\left|\int_a^b f\,dx\right| \le \int_a^b |f|\,dx$$

Note: *The converse of the above theorem is not true*. Consider, for example, the function,

$$f(x) = \begin{cases} 1, & \text{when } x \text{ is rational} \\ -1, & \text{when } x \text{ is irrational} \end{cases}$$

Here

$$\overline{\int_a^b} f\,dx = b - a, \underline{\int_a^b} f\,dx = -(b-a)$$

$\Rightarrow$ f is not integrable

But $|f(x)| - 1$, for all x, so that $\int_a^b |f|dx$ exists and is equal to $(b-a)$.

Thus, $|f|$ is integrable while f is not.

Theorem 11. *If f is integrable on $[a, b]$, then f^2 is also integrable on $[a, b]$.*

Since f is bounded on $[a, b]$, therefore $|f|$ is also bounded on $[a, b]$.

Thus $\exists\ M > 0$, such that $|f(x)| \le M$, for all x in $[a, b]$.

Again, since f is integrable, therefore $|f|$ is also integrable on $[a, b]$, and therefore for $\varepsilon > 0, \exists$ a partition P of $[a, b]$ such that

$$U(P, |f|) - L(P, |f|) < \frac{\varepsilon}{2M}$$

Again, since $|f^2(x)| = |f(x)|^2 \le M^2$, therefore f^2 is bounded.

If M_i, m_i be the bounds of $|f|$ and M_i', m_i' those of f^2 in Δx_i, then $M_i' = M_i^2, m_i' = m_i^2$.

Also

$$\begin{aligned} U(P, f^2) - L(P, f^2) &= \sum_{i=1}^{n} (M_i' - M_i')\Delta x_i \\ &= \sum_{i=1}^{n} (M_i^2 - M_i^2)\Delta x_i \\ &= \sum_{i=1}^{n} (M_i - m_i)(M_1 + m_i)\Delta x_i \\ &\le 2M\,\{\sum_{i=1}^{n} (M_i - m_i)\Delta x_i\} \\ &= 2M\,\{U(P, |f|) - L(P, |f|)\} \\ &< 2M\,\frac{\varepsilon}{2M} = \varepsilon \end{aligned}$$

$\Rightarrow$ $$f^2 \in R\,[a, b]$$

Corollary. If f_1 and f_2 are both integrable on $[a, b]$, then $f_1 f_2$ is also integrable on $[a, b]$.

Since f_1, f_2 are integrable on $[a, b]$, therefore, f_1^2, f_2^2 and $(f_1 + f_2)^2$ are all integrable on $[a, b]$. Also

$$f_1 f_2 = \frac{1}{2}\left\{(f_1 + f_2)^2 - f_1^2 - f_2^2\right\}$$

therefore, $f_1 f_2 \in R\,[a, b]$

6. THE INTEGRAL AS A LIMIT OF SUMS (*Riemann sums*)

Earlier, we arrived at the integral of a function via the upper and the lower sums. The numbers M_i, m_i which appear in these sums are not necessarily the values of the function f (they are values of f if f is continuous). We shall now show that $\int f\,dx$ can also be considered as the limit of a sequence of sums in which M_i and m_i are replaced by the values of f.

Corresponding to a partition P of $[a, b]$, let us choose points $t_1, t_2, \ldots, t_n$ such that $x_{i-1} \le t_i \le x_i$ $(i = 1, 2, \ldots, n)$ and let us consider the sum

$$S(P, F) = \sum_{i=1}^{n} f(t_i)\Delta x_i$$

The sum $S(P, f)$ is called *a Riemann sum of f over* $[a, b]$ *relative to P.*

It may be noted that t_i are arbitrary and that t_i can be any point whatsoever of Δx_i.

We say that $S(P, f)$ converges to A as $\mu(P) \to 0$, *i.e.*,

$$\lim_{\mu(P)\to 0} S(P, f) = A$$

if, for every $\varepsilon > 0$, there exists $\delta > 0$ such that

$$\left|S(P, f) - A\right| < \varepsilon$$

for every partition $P = \{x_0, x_1, \ldots, x_n\}$ of $[a, b]$ with norm $\mu(P) < \delta$ and for every choice of points t_i in $[x_{i-1}, x_i]$.

Definition 2. (*Second definition of integrability*). *A function f is said to be integrable on* $[a, b]$ *if* $\lim S(P, f)$ *exists as* $\mu(P) \to 0$, *and then*

$$\lim_{\mu(P)\to 0} S(P, f) = \int_a^b f\,dx.$$

Note: Since $\mu(P) \to 0$ when $n \to \infty$, therefore $\lim\limits_{\mu(P)\to 0}$ can be replaced by $\lim\limits_{n\to\infty}$ in the above definition.

We have, thus given two definitions of integrability. The equivalence of the two will now be established.

Def. 1 $\Rightarrow$ Def. 2. Let a bounded function f be integrable according to the former definition, so that

$$\int_a^b f\,dx = \overline{\int_a^b} dx = \underline{\int_a^b} f\,dx$$

Let ε be any positive number.

By Darboux's Theorem, there exists $\delta > 0$ such that for every partition P with norm $\mu(P) < \delta$,

$$U(P, f) < \overline{\int_a^b} f\,dx + \varepsilon = \int_a^b f\,dx + \varepsilon \qquad \text{...(1)}$$

and

$$L(P, f) > \underline{\int_a^b} f\,dx - \varepsilon = \int_a^b f\,dx - \varepsilon \qquad \text{...(2)}$$

If t_i is any point of Δx_i, we have

$$L(P, f) \le \sum_{i=1}^{n} f(t_i)\,\Delta x_i \le U(P, f) \qquad \text{...(3)}$$

From equations(1), (2) and (3), we deduce that for any $\varepsilon > 0$, $\exists\, \delta > 0$ such that for every partition P with norm $\mu(P) < \delta$,

$$\int_a^b f\,dx - \varepsilon < \sum_{i=1}^{n} f(t_i)\,\Delta x_i < \int_a^b f\,dx + \varepsilon$$

$$\Rightarrow \qquad \left| \sum_{i=1}^{n} f(t_i)\,\Delta x_i - \int_a^b f\,dx \right| < \varepsilon$$

Thus, the function is integrable according to Def. 2 also.

Remark: Thus $f \in R \Rightarrow \lim S(P, f)$ exists.

Def. 2 $\Rightarrow$ Def. 1. Let a function f be integrable according to the second definition, *i.e.*,

$$\lim_{\mu(P)\to 0} S(P, f) \text{ exists}$$

In other words, to every number $\varepsilon > 0$, there exists $\delta > 0$ such that for every partition $P = \{x_0, x_1, \ldots, x_n\}$ with norm $\mu(P) < \delta$ and for every choice of points t_i in Δx_i, $\exists$ a number A, such that

$$\left| \sum_{i=1}^{n} f(t_i)\,\Delta x_i - A \right| < \varepsilon$$

It will first be shown that f is bounded.

Let, if possible, f be not bounded.

Taking $\varepsilon = 1$, there exists a partition P such that for every choice of t_i in Δx_i

$$\left| \Sigma f(t_i)\Delta x_i - A \right| < 1$$

$$\Rightarrow \qquad \left| \Sigma f(t_i)\Delta x_i \right| < |A| + 1$$

As f is not bounded in $[a, b]$ it must also be so in at least one subinterval, say Δx_m.

Let us take $t_i = x_i$ when $i \neq m$ so that every t_i except t_m is fixed and accordingly every term of the sum $\Sigma f(t_i)\,\Delta x_i$ except the term $f(t_m)\,\Delta x_m$ is also fixed. Since f is not bounded in Δx_m, we can choose a point t_m in Δx_m, such that

$$|\Sigma f(t_i)\Delta x_i| > |A| + 1$$

and thus we arrive at a contradiction.

Hence, the function f is bounded on $[a, b]$.

Now, let ε be any positive number. Thus there exists $\delta > 0$ such that for all partitions P with $\mu(P) < \delta$, we have

$$A - \frac{1}{2}\varepsilon < S(P, f) < A + \frac{1}{2}\varepsilon \qquad ...(4)$$

We choose one such P. If we let the points t_i range over the intervals Δx_i and take the l.u.b. and the g.l.b. of the numbers $S(P, f)$ obtained in this way, (4) yields

$$A - \frac{1}{2}\varepsilon < L(P, f) \leq U(P, f) < A + \frac{1}{2}\varepsilon \qquad ...(5)$$

$$\Rightarrow \qquad U(P, f) - L(P, f) < \varepsilon$$

Also

$$L(P, f) \leq \underline{\int_a^b} f\,dx \leq \overline{\int_a^b} f\,dx \leq U(P, f)$$

$$\therefore \qquad \overline{\int_a^b} f\,dx - \underline{\int_a^b} f\,dx \leq U(P, f) - L(P, f) < \varepsilon$$

so that a non-negative number is less than every positive number.

$$\Rightarrow \qquad \overline{\int_a^b} f\,dx - \underline{\int_a^b} f\,dx = 0$$

or

$$\overline{\int_a^b} f\,dx = \underline{\int_a^b} f\,dx$$

so that, the function is integrable.

6.1 Example 4. Show that $\int_1^2 f\,dx = \dfrac{11}{2}$, where $f(x) = 3x + 1$.

■ Let $P = \{1 = x_0, x_1, x_2, \ldots, x_n = 2\}$ be a partition which divides $[1, 2]$ into n equal sub-intervals, each of length $\dfrac{2-1}{n} = \dfrac{1}{n}$, so that

$$\mu(P) = \frac{1}{n} \to 0 \text{ as } n \to \infty$$

$$x_i = 1 + \frac{i}{n}, \; i = 1, 2, \ldots, n$$

$$\Delta x_i = \frac{1}{n}, i = 1, 2, \ldots, n$$

$$\sum_{i=1}^{n} \Delta x_i = n.\frac{1}{n} = 1$$

Let $t_i = x_i$, when $i = 1, 2, \ldots, n$.

$$\therefore \qquad S(P, f) = \sum_{i=1}^{n} f(t_i)\,\Delta x_i = \sum_{i=1}^{n} f(x_i)\,\Delta x_i = \sum_{i=1}^{n} (3x_i + 1)\,\Delta x_i$$

$$= \sum_{i=1}^{n}\left\{3\left(1+\frac{i}{n}\right)+1\right\}\Delta x_i = 4\sum_{i=1}^{n}\Delta x_i + \frac{3}{n^2}\sum_{i=1}^{n} i$$

$$= 4 + \frac{3}{n^2}\,\frac{n(n+1)}{2} = \frac{11}{2} + \frac{3}{2n}$$

Proceeding to limits when $\mu(P) \to 0$,

$$\lim_{\mu(P)\to 0} S(P, f) = \frac{11}{2}$$

Since, the limit exists, the function is integrable and

$$\int_1^2 f\,dx = \lim S(P, f) = \frac{11}{2}$$

Example 5. Compute $\int_{-1}^{1} f\,dx$, where $f(x) = |x|$.

■ The function f is bounded and continuous on $[-1, 1]$, and

$$f(x) = \begin{cases} -x, & x \le 0 \\ x, & x \ge 0 \end{cases}$$

Let the partition $P = \{-1 = x_0, x_1, \ldots, x_n = 0 = y_0, y_1, \ldots, y_n = 1\}$ divides $[-1, 1]$ into $2n$ equal sub-intervals each of length $1/n$, so that

$$\mu(P) = 1/n \to 0 \text{ as } n \to \infty;$$

$$x_i = -1 + i/n,\ i = 1, 2, \ldots, n$$

$$y_i = i/n,\ i = 1, 2, \ldots, n$$

$$\Delta x_i = 1/n = \Delta y_i$$

$$\sum_{i=1}^{n} \Delta x_i = 1 = \sum_{i=1}^{n} \Delta y_i$$

$$\left.\begin{array}{llll} \text{Let} & t_i \in \Delta x_i & \text{and} & t_i' \in \Delta y_i \\ \text{and let} & t_i = x_i & \text{and} & t_i' = y_i \end{array}\right\}, i = 1, 2, \ldots, n$$

CHAPTER 9

$$
\begin{aligned}
S(P,f) &= \sum_{i=1}^{n} f(t_i)\,\Delta x_i + \sum_{i=1}^{n} f(t_i')\,\Delta y_i \\
&= \sum_i (-x_i)\,\Delta x_i + \sum_i y_i\,\Delta y_i \\
&= \sum_i \left(1 - \frac{i}{n}\right)\Delta x_i + \sum_i \frac{i}{n}\Delta y_i \\
&= \sum_{i=1}^{n} \Delta x_i - \frac{1}{n^2}\sum_{i=1}^{n} i + \frac{i}{n^2}\sum_{i=1}^{n} i
\end{aligned}
$$

$$\therefore \qquad \lim_{\mu(P)\to 0} S(P,f) = 1$$

and since the limit exists, the function is integrable and

$$\int_{-1}^{-b} |x|\,dx = \lim S(P,f) = 1.$$

6.2 Some Applications

I. *If f_1 and f_2 are functions and where $f = f_1 \pm f_2$ are bounded and integrable on* $[a, b]$, *then*

$$\int_a^b f\,dx = \int_a^b f_1\,dx \pm \int_a^b f_2\,dx$$

Let ε be any positive number and $f = f_1 + f_2$.

Since f_1, f_2 are integrable, therefore $\exists\ \delta > 0$ such that for every partition $P = \{x_0, x_1, x_2, \ldots, x_n\}$ with norm $\mu(P) < \delta$ and for every choice of points t_i in Δx_i.

$$\left.\begin{aligned}
\left|\sum_i f_1(t_i)\Delta x_i - \int_a^b f_1\,dx\right| < \frac{1}{2}\varepsilon \\
\left|\sum_i f_2(t_i)\Delta x_i - \int_a^b f_2\,dx\right| < \frac{1}{2}\varepsilon
\end{aligned}\right\}$$

$$\Rightarrow \qquad \left|\sum \{(f_1 + f_2)(t_i)\}\,\Delta x_i - \left\{\int_a^b f_1\,dx + \int_a^b f_2\,dx\right\}\right| < \varepsilon$$

$$\therefore \qquad \int_a^b f\,dx = \int_a^b (f_1 + f_2)\,dx = \int_a^b f_1\,dx + \int_a^b f_2\,dx$$

The case of $f = f_1 - f_2$ may be discussed similarly.

Corollary. *If* $f_1 \in R$ and $f_2 \in R$ *over* $[a, b]$, *and* c_1, c_2 *any two constants, then* $c_1 f_1 + c_2 f_2 \in R$ $[a, b]$ *and*

$$\int_a^b (c_1 f_1 + c_2 f_2)\, dx = c_1 \int_a^b f_1\, dx + c_2 \int_a^b f_2\, dx$$

Let $f = c_1 f_1 + c_2 f_2$.

For any partition P, we can write

$$S(P, f) = \sum_i f(t_i)\, \Delta x_i = c_1 \sum_i f_1(t_i)\, \Delta x_i + c_2 \sum_i f_2(t_i)\, \Delta x_i$$

$$= c_1 S(P, f_1) + c_2 S(P, f_2)$$

Since f_1, f_2 are integrable, for $\varepsilon > 0$, we can choose $\delta > 0$ such that for all partition P with $\mu(P) < \delta$, we have

$$\left| S(P, f_1) - \int_a^b f_1\, dx \right| < \varepsilon$$

and
$$\left| S(P, f_2) - \int_a^b f_2\, dx \right| < \varepsilon$$

$\therefore$
$$\left| S(P, f) - c_1 \int_a^b f_1\, dx - c_2 \int_a^b f_2\, dx \right| < c_1\, \varepsilon + c_2\, \varepsilon$$

$\Rightarrow$ $\lim\limits_{\mu(P) \to 0} S(P, f)$ exists and equals $c_1 \int_a^b f_1\, dx + c_2 \int_a^b f_2\, dx$

Hence, $(c_1 f_1 + c_2 f_2) \in R$ and

$$\int_a^b (c_1 f_1 + c_2 f_2)\, dx = c_1 \int_a^b f_1\, dx + c_2 \int_a^b f_2\, dx$$

II. *If a function f is bounded and integrable on each of the inter vals* $[a, c]$, $[c, b]$, $[a, b]$ *where c is a point of* $[a, b]$, *then*

$$\int_a^b f\, dx = \int_a^c f\, dx + \int_c^b f\, dx$$

Let $\varepsilon > 0$ be given.

As f is integrable on each of the intervals $[a, c]$, $[c, b]$ and $[a, b]$, there exists $\delta > 0$ such that for every partition $P = \{a = x_0, x_1, \ldots, x_n = b\}$ containing the point c, with norm $\mu(P) < \delta$ and for every choice of points t_i in Δx_i,

$$\left| \sum_{[a,c]} f(t_i) \Delta x_i - \int_a^c f\, dx \right| < \frac{1}{3}\varepsilon$$

CHAPTER 9

$$\left| \sum_{[c,b]} f(t_i)\Delta x_i - \int_c^b f\,dx \right| < \frac{1}{3}\varepsilon$$

$$\left| \sum_{[a,b]} f(t_i)\Delta x_i - \int_a^b f\,dx \right| < \frac{1}{3}\varepsilon$$

But

$$\sum_{[a,c]} f(t_i)\,\Delta x_i + \sum_{[c,b]} f(t_i)\,\Delta x_i = \sum_{[a,b]} f(t_i)\,\Delta x_i$$

Therefore, we deduce that

$$\left| \int_a^b f\,dx - \int_a^c f\,dx - \int_c^b f\,dx \right| < \varepsilon$$

$$\Rightarrow \qquad \int_a^b f\,dx = \int_a^c f\,dx + \int_c^b f\,dx$$

III. *A function f is integrable over $[a, b]$ iff for $\varepsilon > 0, \exists\, \delta > 0$ such that if P, P' are any two partitions of $[a, b]$ with mesh less than δ, then*

$$|S(P,f) - S(P',f)| < \varepsilon$$

[This is the analog for Riemann sums of the Cauchy property of sequences.]

First, let $f \in R\,[a, b]$, and $\int_a^b f\,dx = I$

$\therefore$ For $\varepsilon > 0, \exists\, \delta > 0$ such that for all partitions P, P' of $[a, b]$ with mesh less than δ and all positions of t_i in Δx_i,

$$|S(P,f) - I| < \frac{1}{2}\varepsilon \qquad \text{...(1)}$$

$$|S(P',f) - I| < \frac{1}{2}\varepsilon \qquad \text{...(2)}$$

$$\Rightarrow |S(P,f) - S(P',f)| \le |S(P,f) - I| + |S(P',f) - I| < \frac{1}{2}\varepsilon + \frac{1}{2}\varepsilon = \varepsilon$$

Conversely, for $\varepsilon > 0\ \exists\, \delta_1 > 0$ such that for any partitions P, P' with less than δ_1 we have

$$|S(P,f) - S(P',f)| < \frac{1}{2}\varepsilon \qquad \text{...(3)}$$

We know that for a given partition P' and every choice of t_i in Δx_i, $S(P', f)$ is bounded by $L(P', f)$ and $U(P', f)$, which for all partitions of $[a, b]$ are, in turn, bounded by $m\,(b - a)$ and $M\,(b - a)$, where m, M are bounds of f.

Thus, the sequence $\{S(P', f)\}$ of Riemann sums is bounded. As every bounded sequence has a limit point, let the sequence have a limit point I, mesh so that

$$\lim_{\mu(P')\to 0} S(P', f) = I$$

Hence, for $\varepsilon > 0, \exists\, \delta_2 > 0$ such that for partition P' with mesh

$$\mu(P') < \delta_2$$

$$\left|S(P',f) - I\right| < \frac{1}{2}\varepsilon \qquad \text{...(4)}$$

Let $\delta = \min(\delta_1, \delta_2)$.

$\therefore$
$$\left|S(P,f) - I\right| \le \left|S(P,f) - S(P',f)\right| + \left|S(P',f) - I\right|$$
$$< \frac{1}{2}\varepsilon + \frac{1}{2}\varepsilon = \varepsilon, \text{ for } \mu(P) < \delta$$

$\Rightarrow$
$$\lim_{\mu(P)\to 0} S(P,f) = I$$

Thus, $\lim_{\mu(P)\to 0} S(P,f)$ exists and hence, the function f is integrable.

7. SOME INTEGRABLE FUNCTIONS

Theorem 12. *Every continuous function is integrable.*

We shall prove that a function f which is continuous on $[a, b]$ is also integrable on $[a, b]$.

Let $\varepsilon > 0$ be given.

Let us choose a positive number η, such that

$$\eta(b - a) < \varepsilon$$

Since f is continuous on the closed interval $[a, b]$, therefore, it is bounded and is *uniformly continuous* on $[a, b]$, which implies that there exists $\delta > 0$, such that

$$\left|f(t_1) - f(t_2)\right| < \eta, \text{ if } \left|t_1 - t_2\right| < \delta, \text{ and } t_1, t_2 \in [a, b] \qquad \text{...(1)}$$

We, now, choose a partition P with norm $\mu(P) < \delta$.

Then by equation (1), we have

$$M_i - m_i \le \eta \quad (i = 1, 2, \ldots, n)$$

$\therefore$
$$U(P,f) - L(P,f) = \sum_{i=1}^{n} (M_i - m_i)\, \Delta x_i$$
$$\le \eta \sum_{i=1}^{n} \Delta x_i$$
$$= \eta(b - a) < \varepsilon \qquad \text{...(2)}$$

Thus, f is integrable [§ 4 Theorem 3].

Corollary. *If a function f is continuous, then to every $\varepsilon > 0$ there corresponds $\delta > 0$ such that*

$$\left|\sum_{i=1}^{n} f(t_i)\, \Delta x_i - \int_a^b f\, dx\right| < \varepsilon$$

for every partition $P = \{x_0, x_1, x_2, \ldots, x_n\}$ of $[a, b]$ with $\mu(P) < \delta$, and for every choice of points t_i in $[x_{i-1}, x_i]$.

Follows from (2) above, since the two numbers $\sum f(t_i)\, \Delta x_i$ and $\int f\, dx$ lie between $U(P, f)$ and $L(P, f)$.

Remark: For continuous functions, $\lim_{\mu(P)\to 0} \sum f(t_i)\,\Delta x_i$ exists and equals

$$\int_a^b f\,dx.$$

Note: Continuity is a sufficient condition for integrability. It is not a necessary condition. Functions exist which are integrable but not continuous. See examples 6, 7 and 8. Many more may be constructed.

Theorem 13. *If a function f is monotonic on* $[a, b]$, *then it is integrable on* $[a, b]$.

We shall prove the theorem when f is monotonic increasing (the proof for the other case is analogous). Clearly f is bounded.

Let $\varepsilon > 0$ be given.

Let us choose a number $\eta < \dfrac{\varepsilon}{f(b) - f(a) + 1}$.

We, now, choose a partition $P = \{x_0, x_1, \ldots, x_n\}$ of $[a, b]$ with norm $\mu(P) < \eta$.

Since f is monotonic increasing, therefore

$$M_i = f(x_i),\ m_i = f(x_{i-1}), \quad (i = 1, 2, \ldots, n)$$

$$\begin{aligned}
U(P,f) - L(P,f) &= \sum_{i=1}^{n} (M_i - m_i)\,\Delta x_i \\
&= \sum_{i=1}^{n} \{f(x_i) - f(x_{i-1})\}\,\Delta x_i \\
&< \eta \sum_{i=1}^{n} \{f(x_i) - f(x_{i-1})\} \\
&< \frac{\varepsilon}{f(b) - f(a) + 1} \cdot \{f(b) - f(a)\} \\
&< \varepsilon.
\end{aligned}$$

Hence, f is integrable.

Note: $f(b) - f(a) + 1$ have been taken to cover the case when $f(a) = f(b)$.

Theorem 14. *A bounded function f, having a finite number of points of discontinuity on* $[a, b]$, *is integrable on* $[a, b]$.

Let M, m be the bounds of f.

Let $\varepsilon > 0$ be a given number.

Let there be p points of discontinuity of f on $[a, b]$.

We now consider a partition P of $[a, b]$ such that all the points of discontinuity get enclosed in p non-overlapping sub-intervals, the sum of whose lengths $< \frac{1}{2}\varepsilon/(M - m)$. The oscillation of f in each of these sub-intervals being $\le (M - m)$, their total contribution to the difference $\{U(P,f) - L(P,f)\}$ is less than

$$\frac{\varepsilon}{2(M-m)}(M-m)=\frac{1}{2}\varepsilon.$$

The function f is continuous in the remaining portion of $[a, b]$, *i.e.*, in the $(p + 1)$ sub-intervals of $[a, b]$ excluding the sub-intervals considered above.

As in Theorem 12, the contribution to the difference $\{U(P, f) — L(P, f)\}$ from each of these $(p + 1)$ sub-intervals can be made $< \dfrac{\varepsilon}{2(p+1)}$, so that the total contribution to $\{U(P, f) — L(P, f)\}$ by these $(p + 1)$ sub-intervals is less than $\dfrac{\varepsilon}{2(p+1)}(p+1)=\dfrac{1}{2}\varepsilon$

Thus, for the partition P of $[a, b]$,

$$U(P,f)-L(P,f)<\varepsilon.$$

Hence, the function f is integrable.

Theorem 15. *A bounded function f is integrable on* $[a, b]$, *if the set of its points of discontinuity has only a finite number of limit points.*

Let the set of points of discontinuity of f have a finite number p of limit points. Let M, m be the bounds of f.

The limit points may be enclosed in p non-overlapping subintervals of $[a, b]$, the sum of whose lengths $\leq \frac{1}{2}\varepsilon / (M-m)$. So that their total contribution to $\{U(P, f) - L(P, f)\}$ is $< \frac{1}{2}\varepsilon$

Only a finite number of points of discontinuity of f can be outside these subintervals, *i.e.*, the function f has a finite number of points of discontinuity on $[a, b]$ excluding the p subintervals enclosing the limit points. Therefore, as in theorem 14, the total contribution to $\{U(P,f)-L(P,f)\}$ from these portions of $[a, b]$ can be made $< \frac{1}{2}\varepsilon$.

Thus for such a partition P of $[a, b]$,

$$U(P,f)-L(P,f)<\varepsilon.$$

Hence, the function f is integrable on $[a, b]$.

Example 6. A function f is defined on $[-1, 1]$ as follows:

$$f(x)=\begin{cases} k, \text{ positive constant when } x\neq 0 \\ 0, \text{ when } x=0. \end{cases}$$

Show that f is integrable on $[-1, 1]$ and that the value of the integral is $2k$.

■ The function has only one point of discontinuity, 0, and is therefore integrable (Theorem 14). Proceeding as in § 1.1 Example 1, it may be shown that the value of the integral is $2k$.

Example 7. A function f is defined on $[0, 1]$ as follows:

$$f(x)=\begin{cases} 0, \text{ when } x \text{ is irrational or zero, and} \\ 1/q, \text{ when } x \text{ is any non-zero rational number } p/q \text{ with least positive integers } p \text{ and } q. \end{cases}$$

Show that f is integrable on $[0, 1]$ and the value of the integral is zero.

■ The function f is bounded with bounds 0 and 1.

Let ε be any positive number.

There exists the largest integer $q \in \mathbf{N}$ such that $1/q > \frac{1}{2}\varepsilon$ or $q < 2/\varepsilon$, so that there are only a finite number of points p/q for which $I/q > \frac{1}{2}\varepsilon$. Let us call such points as exceptional points.

Thus at those rational points which are exceptional points, f has a value $1/q > \frac{1}{2}\varepsilon$, while at the other rational points, the value of f is $1/q > \frac{1}{2}\varepsilon$. The function is zero at the irrational points.

Also, every interval contains rational as well as irrational points.

Thus the oscillation of f in any interval which includes no exceptional points is less than $\frac{1}{2}\varepsilon$ and that in an interval which includes the exceptional points it is at the most equal to 1.

Let us consider a partition P of [0, 1] so as to enclose the exceptional points (finite in number) in subintervals, the sum of whose lengths is less than $\frac{1}{2}\varepsilon$. Thus the contribution to $[U(P,f) - L(P,f)]$ made by these is less than $\frac{1}{2}\varepsilon$.

The contribution to $[U(P,f) - L(P,f)]$ by the remaining portion of [0, 1] is evidently less than $\frac{1}{2}\varepsilon$.

Hence, $U(P,f) - L(P,f) < \varepsilon$, so that the function f is integrable.

Note: The above function has a discontinuity at each *rational*, while continuity at each *irrational*.

Example 8. Show that the function f defined as follows:

$$f(x) = \frac{1}{2^n}, \text{ when } \frac{1}{2^{n+1}} < x \leq \frac{1}{2^n}, \ (n = 0, 1, 2, \ldots),$$

$$f(0) = 0,$$

is integrable on [0, 1], although it has an infinite number of points of discontinuity.

■ Now

$$\begin{aligned} f(x) &= 1, \text{ when } \frac{1}{2} < x \leq \frac{1}{1} \\ &= \frac{1}{2}, \text{ when } \frac{1}{2^2} < x \leq \frac{1}{2} \\ &= \frac{1}{2^2}, \text{ when } \frac{1}{2^3} < x \leq \frac{1}{2^2} \\ &\ \ \vdots \\ &= \frac{1}{2^{n-1}}, \text{ when } \frac{1}{2^n} < x \leq \frac{1}{2^{n-1}} \\ &\ \ \vdots \\ &= 0, \text{ when } x = 0. \end{aligned}$$

Thus, we notice that f is bounded and monotonic increasing on [0, 1].
Hence, f is integrable [Theorem13].

Aliter. f is continuous on $[0, 1]$, except at the set of points

$$0, \frac{1}{2}, \left(\frac{1}{2}\right)^2, \left(\frac{1}{2}\right)^3, \ldots$$

which have only one limit point, 0, and hence f is integrable.

8. INTEGRATION AND DIFFERENTIATION (*The Primitive*)

We shall first show that integration and differentiation are in a certain sense, inverse operations, then define the primitive of a function and go on to prove a theorem which is usually called the fundamental theorem of calculus.

Theorem 16. *If a function f is bounded and integrable on $[a, b]$, then the function F defined as*

$$F(x) = \int_a^x f(t)\,dt,\ a \le x \le b$$

is continuous on $[a, b]$, and furthermore, if f is continuous at a point c of $[a, b]$, then F is derivable at c and

$$F'(c) = f(c).$$

Since f is bounded therefore $\exists$ a number $K > 0$, such that

$$|f(x)| \le K,\quad \forall\, x \in [a, b].$$

If x_1, x_2 are two points of $[a, b]$ such that $a \le x_1 < x_2 \le b$, then

$$\begin{aligned} |F(x_2) - F(x_1)| &= \left| \int_{x_1}^{x_2} f(t)\,dt \right| \\ &\le K(x_2 - x_1) \qquad (\S\ 1.3 \text{ Deduction } 3) \end{aligned}$$

Thus for a given $\varepsilon > 0$, we see that

$$|F(x_2) - F(x_1)| < \varepsilon,\ \text{ if } |x_2 - x_1| < \varepsilon/K.$$

Hence, the function F is continuous (in fact uniformly) on $[a, b]$.

Let f be continuous at a point c of $[a, b]$, so that for any $\varepsilon > 0$ there exists $\delta > 0$ such that

$$|f(x) - f(c)| < \varepsilon,\ \text{ for } |x - c| < \delta$$

Let $c - \delta < s \le c \le t < c + \delta$

$$\begin{aligned} \therefore \qquad \left| \frac{F(t) - F(s)}{t - s} - f(c) \right| &= \left| \frac{1}{t - s} \int_s^t \{f(x) - f(c)\}\,dx \right| \\ &\le \frac{1}{t - s} \int_s^t |f(x) - f(c)|\,dx < \varepsilon. \end{aligned}$$

$$\Rightarrow \qquad F'(c) = f(c),$$

i.e., continuity of f at any point of $[a, b]$ implies derivability of **F** at that point.

CHAPTER 9

Note: As c is any point of $[a, b]$, we have for all $x \in [a, b]$,

$$F'(t) = f(t) \qquad \Rightarrow F' = f$$

i.e., continuity of f on $[a, b]$ implies derivability of F on $[a, b]$.

This theorem is sometimes referred to as the *First Fundamental Theorem of Integral Calculus*.

Definition. A derivable function F, if it exists such that its derivative F' is equal to a given function f, is called a *primitive* of f.

The above theorem shows that a sufficient condition for a function to admit of a primitive is that it is continuous. Thus every continuous function f possesses a primitive F, where

$$F(x) = \int_a^x f(t)\, dt$$

Remark: We shall now show, with the help of an example, that continuity of a function is not a necessary condition for the existence of a primitive, in other words, "*functions possessing primitives are not necessarily continuous*". Consider the function f on $[0, 1]$, where

$$f(x) = \begin{cases} 2x, \sin(1/x) - \cos(1/x), & \text{if } x \neq 0 \\ 0, & \text{if } x = 0. \end{cases}$$

It has a primitive F, where

$$F(x) = \begin{cases} x^2 \sin 1/x, & \text{if } x \neq 0. \\ 0, & \text{if } x = 0. \end{cases}$$

Clearly $F'(x) = f(x)$ but $f(x)$ is not continuous at $x = 0$, *i.e.*, $f(x)$ is not continuous on $[0, 1]$ [Example 4, Ch. 6]. In fact, all this amounts to saying that the derivative of a function is not necessarily continuous.

9. THE FUNDAMENTAL THEOREM OF CALCULUS

Theorem 17. *A function f is bounded and integrable on $[a, b]$, and there exists a function F such that $F' = f$ on $[a, b]$, then*

$$\int_a^b f\, dx = F(b) - F(a).$$

Since the function $F' = f$ is bounded and integrable, therefore for every given $\varepsilon > 0, \exists\, \delta > 0$ such that for every partition $P = \{a = x_0, x_1, x_2, \ldots, x_n = b\}$, with norm $\mu(P) < \delta$,

$$\left.\begin{aligned} &\left| \sum_{i=1}^{n} f(t_i)\, \Delta x_i - \int_a^b f\, dx \right| < \varepsilon \\ &\text{or} \\ &\lim_{(P)\to 0} \sum_{i=1}^{n} f(t_i)\, \Delta x_i = \int_a^b f\, dx, \end{aligned}\right\} \qquad \ldots(1)$$

for every choice of points t_i in Δx_i.

Since we have freedom in the selection of points t_i in Δx_i, we choose them in a particular way as follows:

By Lagrange's Mean Value Theorem, we have

$$F(x_i) - F(x_{i-1}) = F'(t_i)\,\Delta x_i \quad (i = 1, 2, \ldots, n)$$

$$= f(t_i)\,\Delta x_i$$

$$\Rightarrow \qquad \sum_{i=1}^{n} f(t_i)\,\Delta x_i = \sum_{i=1}^{n} \{F(x_i) - F(x_{i-1})\}$$

$$= F(b) - F(a).$$

Hence, from (1) $\qquad \int_a^b f\,dx = F(b) - F(a)$

It is sometimes referred to as *The Second Fundamental Theorem of Integral Calculus*.

9.1 Solved Examples

Example 9. Show that $\int_0^1 f\,dx = \frac{2}{3}$, where f is the integrable function in Example 8.

■ The function f is integrable

$$\therefore \qquad \int_{1/2^n}^{1} f\,dx = \int_{1/2}^{1} f\,dx + \int_{1/2^2}^{1/2} f\,dx + \int_{1/2^3}^{1/2^2} f\,dx + \ldots + \int_{1/2^n}^{1/2^{n-1}} f\,dx$$

$$= \int_{1/2}^{1} dx + \frac{1}{2}\int_{1/2^2}^{1/2} dx + \frac{1}{2^2}\int_{1/2^3}^{1/2^2} dx + \ldots + \frac{1}{2^{n-1}}\int_{1/2^n}^{1/2^{n-1}} dx$$

$$= \frac{1}{2} + \frac{1}{2}\left(\frac{1}{2} - \frac{1}{2^2}\right) + \frac{1}{2^2}\left(\frac{1}{2^2} - \frac{1}{2^3}\right) + \ldots + \frac{1}{2^{n-1}}\left(\frac{1}{2^{n-1}} - \frac{1}{2^n}\right)$$

$$= \frac{1}{2}\left\{1 + \frac{1}{2^2} + \left(\frac{1}{2^2}\right)^2 + \left(\frac{1}{2^2}\right)^3 + \ldots + \left(\frac{1}{2^2}\right)^{n-1}\right\}$$

$$= \frac{1}{2}.\frac{1-\frac{1}{4}}{1-\frac{1}{4}} = \frac{2}{3}\left(1 - \frac{1}{4^n}\right)$$

Proceeding to limits when $n \to \infty$, we get

$$\int_0^1 f\,dx = \frac{2}{3}$$

Example 10. Show that the function $[x]$, where $[x]$ denotes the greatest integer not greater than x, is integrable in $[0, 3]$, and

$$\int_0^3 [x]\,dx = 3$$

CHAPTER 9

- The function is bounded and has only three points of finite discontinuity at 1, 2, 3.

  Let ε be a given positive number.

  Consider a partition P where,

$$P = \{0 = x_0, x_1, \ldots, x_l, y_0, y_1, \ldots, y_m, z_0, z_1, \ldots, z_n = 3\}$$

where $y_0 = 1,\ z_0 = 2$.

Now

$$\begin{aligned} U(P, f) &= \Sigma\, 0 \cdot \Delta x_i + 1 \cdot (y_0 - x_l) + \Sigma\, 1 \cdot \Delta y_i + 2(z_0 - y_m) \\ &\qquad + \Sigma\, 2 \cdot \Delta z_i + 3(z_n - z_{n-1}) \\ &= 1 + 2 + \{(y_0 - x_l) + (z_0 - y_m) + (z_n - z_{n-1})\} \end{aligned}$$

Let us select P such that

$$(y_0 - x_i) + (z_0 - y_m) + (z_n - z_{n-1}) < \varepsilon$$

$\therefore$ $$U(P, f) < 3 + \varepsilon$$

Again,

$$L(P, f) = \Sigma 0 \cdot \Delta x_i + \Sigma 1 \cdot \Delta y_i + (z_0 - y_m) + \Sigma 2 \cdot \Delta z_i = 3$$

$\therefore$ $$U(P, f) - L(P, f) < \varepsilon$$

so that the function is integrable, and therefore

$$\overline{\int_0^3} [x]\, dx = \underline{\int_0^3} [x]\, dx = 3$$

Aliter. Since the function is bounded and has only three points of discontinuity therefore it is integrable, and

$$\begin{aligned} \int_0^3 [x]\, dx &= \int_0^1 [x]\, dx + \int_1^2 [x]\, dx + \int_2^3 [x]\, dx \\ &= \int_0^1 0\, dx + \int_1^2 dx + \int_2^3 2\, dx = 3 \end{aligned}$$

1, 2 and 3 being respectively the points of discontinuity of the three integrals on the right.

Example 11. Let f is a non-negative continuous function on $[a, b]$ and

$$\int_a^b f\, dx = 0. \text{ Prove that } f(x) = 0, \text{ for all } x \in [a, b].$$

- Suppose that, for some $c \in]a, b[$, $f(c) > 0$.

  Then, for $\varepsilon = \dfrac{1}{2} f(c) > 0$, continuity of f at c implies that, there exists a $\delta > 0$ such that

$$f(x) > \frac{1}{2} f(c), \quad \forall\, x \in]c - \delta, c + \delta[$$

Now,

$$\int_a^b f(x)\,dx = \int_a^{c-\delta} f(x)\,dx + \int_{c-\delta}^{c+\delta} f(x)\,dx + \int_{c+\delta}^{b} f(x)\,dx$$

$$\geq \int_{c-\delta}^{c+\delta} f(x)\,dx \qquad (\because f(x) \geq 0, \;\; \forall\, x \in [a, b])$$

$$> \frac{1}{2} f(c) \int_{c-\delta}^{c+\delta} dx = \delta\, f(x) > 0.$$

which is a contradiction. Thus $f(x) = 0, \;\; \forall\, x \in \,]a, b[$.

Similarly, $f(a) \ngtr 0$, and $f(b) \ngtr 0$. Hence, the result follows.

Example 12. Show that $\int_0^t \sin x\, dx = 1 - \cos t$.

■ The function $\sin x$ is bounded and continuous in any interval $[0, t]$ and is therefore integrable. [To be more specific, take $t \leq \pi/2$.]

Consider a partition $P = \left\{0, \frac{t}{n}, \frac{2t}{n}, \ldots, \frac{nt}{n}\right\}$ of $[0, t]$.

Now

$$U(P, \sin x) = \frac{t}{n}\left\{\sin\frac{t}{n} + \sin\frac{2t}{n} + \sin\frac{3t}{n} + \ldots + \sin\frac{nt}{n}\right\}$$

$$= \frac{t}{n}\left(\cos\frac{t}{2n} - \cos\,(n+1)\,\frac{t}{n}\right)\Big/\left(2\sin\frac{t}{2n}\right)$$

$$= \left\{\cos\frac{t}{2n} - \cos\,(n+1)\,\frac{t}{n}\right\}\Big/\left(\frac{\sin t/2n}{t/2n}\right)$$

and

$$L(P, \sin x) = \frac{t}{n}\left\{\sin\frac{t}{n} + \sin\frac{2t}{n} + \ldots + \sin\overline{n-1}\,\frac{t}{n}\right\}$$

$$= \left\{\cos\frac{t}{2n} - \cos\frac{nt}{n}\right\}\Big/\left(\frac{\sin t/2n}{t/2n}\right).$$

In the limit,

$$U(P, \sin x) = L(P, \sin x) = 1 - \cos t$$

$$\therefore \qquad \int_0^t \sin x\, dx = 1 - \cos t$$

CHAPTER 9

Example 13. Evaluate $\int_0^t \left(2x \sin \frac{1}{x} - \cos \frac{1}{x}\right) dx$

- The function

$$f(x) = \begin{cases} 2x \sin \frac{1}{x} - \cos \frac{1}{x}, & x \in]0, 1[\\ 0, & x = 0 \end{cases}$$

is not continuous on [0, 1] (it is discontinuous at $x = 0$), but it is bounded and continuous on] 0, 1] and thus Riemann-integrable on [0, 1].

The function

$$g(x) = \begin{cases} x^2 \sin \frac{1}{x}, & x \in]0, 1] \\ 0, & x = 0 \end{cases}$$

is differentiable on [0, 1] and satisfies

$$g'(x) = f(x), \quad \forall\, x \in [0, 1]$$

$$\therefore \qquad \int_0^1 \left(2x \sin \frac{1}{x} - \cos \frac{1}{x}\right) dx = g(1) - g(0) = \sin 1$$

Note: If f is not integrable but $f(x) = g'(x), \ \forall\, x \in [a, b]$, then

$$\int_a^b f\, dx \neq g(b) - g(a).$$

This can be seen by the following example, which has a primitive without being Riemann-integrable.

Example 14. Prove with the help of an example that the equation (1)

$$\int_a^b f'(x)\, dx = f(b) - f(a), \text{ is not always valid.}$$

- Let f be defined on [0, 1] as follows:

$$f(x) = x^2 \cos(\pi/x^2), \text{ if } 0 < x \leq 1, f(0) = 0.$$

Then f is differentiable on [0, 1] and

$$f'(x) = 2x \cos(\pi/x^2) + (2\pi/x) \sin(\pi/x^2), \text{ if } 0 < x \leq 1,$$

$$f'(0) = 0$$

f' is not bounded on [0, 1] and therefore, it is not Riemann-integrable, *i.e.*, $\int_0^1 f'(x)\, dx$ does not exist.

Therefore the equation (1) fails to hold.

Note: In fact, there are functions f with bounded derivatives f' that are not Riemann-integrable, but these are much more difficult to construct.

Ex. Find the error in the following:

(*i*) If $f(x) = \dfrac{-1}{x-1}$, $f'(x) = \dfrac{1}{(x-1)^2}$, hence

$$\int_0^2 \frac{1}{(x-1)^2}\,dx = f(2) - f(0) = \frac{-1}{2-1} - \left(\frac{-1}{0-1}\right) = -2$$

(*ii*) If $f(x) = 2\sqrt{x}$, then $f'(x) = \dfrac{1}{\sqrt{x}}$; hence

$$\int_0^1 \frac{1}{\sqrt{x}}\,dx = 2\sqrt{1} - 0 = 2.$$

Note: In both the parts $f'(x)$ is not bounded.

10. MEAN VALUE THEOREMS OF INTEGRAL CALCULUS

We have two mean value theorems for derivatives. In the same way we have two mean value theorems for integrals as well.

10.1 First Mean Value Theorem

Theorem 18. *If a function f is continuous on* $[a, b]$, *then there exists a number* ξ *in* $[a, b]$ *such that*

$$\int_a^b f\,dx = f(\xi)(b-a)$$

f is continuous, therefore $f \in R[a, b]$.

Let m, M be the infimum and supremum of f in $[a, b]$. Then as in § 1.1, we have

$$m(b-a) \le \int_a^b f\,dx \le M(b-a)$$

Hence, there is a number $\mu \in [m, M]$, such that

$$\int_a^b f\,dx = \mu\,(b-a)$$

Since f is continuous on $[a, b]$, it attains every value between its bounds m, M. Therefore, there exists a number $\xi \in [a, b]$ such that $f(\xi) = \mu$.

$$\therefore \qquad \int_a^b f\,dx = f(\xi)(b-a).$$

Remark: The preceding theorem says that the condition of continuity is necessary for the function to assume its mean value in the given interval, for example the function $f(x)$ is defined on [2, 5] as follows:

$$f(x) = \begin{cases} 1 & \text{if } 2 \le x < 3 \\ 3 & \text{if } 3 \le x \le 5 \end{cases}$$

Now

$$\int_2^5 f\,dx + \int_2^3 1\,dx + \int_3^5 3\,dx = 7$$

and so the mean value of the function is

$$\frac{1}{5-2}\int_2^5 f\,dx = \frac{7}{3},$$

which the function fails to assume on the interval.

10.2 The Generalised First Mean Value Theorem

Theorem 19. *If f and g are integrable on [a, b] and g keeps the same sign over [a, b], then there exists a number μ lying between the bounds of f such that*

$$\int_a^b fg\,dx = \mu \int_a^b g\,dx$$

Let g be positive over $[a, b]$.

If m, M are the bounds of f, we have for all $x \in |a, b|$,

$$m \le f(x) \le M$$

$$\Rightarrow \qquad mg(x) \le f(x)\,g(x) \le Mg(x)$$

$$\therefore \qquad m\int_a^b g(x)\,dx \le \int_a^b f(x)\,g(x)\,dx \le M\int_a^b g(x)\,dx, \text{ when } b \ge a \qquad \ldots(1)$$

Let μ be a number lying between m and M.

$$\therefore \qquad \int_a^b fg\,dx = \int_a^b g\,dx \qquad \ldots(2)$$

Corollary. If in addition to the conditions of the above theorem, f is continuous on $[a, b]$, then $\exists$ a number $\xi \in [a, b]$, such that

$$\int_a^b fg\,dx = f(\xi)\int_a^b g\,dx \qquad \ldots(3)$$

Notes:

1. If $g(x) \le 0$, the sign of the inequality changes in (1) but (2) and (3) remain unchanged.
2. If $b \le a$, the sign of inequality changes in (1) but (2) and (3) remain unaffected.

EXERCISE

1. Show that $\int_0^1 x^4\, dx = \frac{1}{5}$.

2. If f is continuous and non-negative on $[a, b]$. Then show that

$$\int_a^b f\, dx \geq 0.$$

3. Determine whether f is Riemann-integrable on [0, 1] and justify your answer.

(*i*) $f(x) = \dfrac{1}{x+1}$

(*ii*) $f(x) = \left|x - \dfrac{1}{4}\right|$

(*iii*) $$\begin{aligned} f(x) &= x \cos \frac{1}{x}, & x \neq 0 \\ &= 0, & x = 0 \end{aligned}$$

(*iv*) $$\begin{aligned} f(x) &= \frac{1}{x-1}, & x \neq 1 \\ &= 0, & x = 1 \end{aligned}$$

(*v*) $$\begin{aligned} f(x) &= \sin \frac{1}{x}, & \text{if } x \text{ is irrational} \\ &= 0, & \text{otherwise} \end{aligned}$$

[*Hint* (*v*): Let P be any partition of [0, 1], then $L(P, f) = 0$. Suppose $\frac{2}{\pi} \in [x_{i-1}, x_i]$. If $\frac{2}{\pi} \leq x \leq 1$, then $\sin \frac{1}{x} \geq \sin 1$, and hence

$$U(P, f) \geq (\sin 1)\,(1 - x_{i-1}) \geq (\sin 1)\,(1 - 2/\pi)$$

It follows that

$$\underline{\int} f\, dx = 0 \text{ and } \overline{\int} f\, dx \geq (\sin 1)\left(1 - \frac{2}{\pi}\right) > 0.$$

Thus, f is not Riemann-integrable.

4. Prove that the function f defined on [0, 1] as

$$\begin{aligned} f(x) &= 2n, \text{ if } x = \frac{1}{n} \text{ where } n = 1, 2, \ldots \\ &= 0, \text{ otherwise} \end{aligned}$$

is not Riemann-integrable on [0, 1].

(*Hint*: $\lim\limits_{x \to 0} f(x) = \infty$, f is not bounded above and so not Riemann integrable on [0, 1])

CHAPTER 9

5. Prove that the function f defined as

$$f(x) = \begin{cases} x, & \text{when } x \text{ is rational} \\ -x, & \text{when } x \text{ is not rational} \end{cases}$$

is not integrable over $[a, b]$; but $|f|$ is integrable.

6. Integrate on $[0, 2]$ the function $f(x) = x[x]$, where $[x]$ denotes the greatest integer not greater than x.

7. f and g are two bounded functions on $[a, b]$ such that $f(x) = g(x)$ except for a finite number of points x in $[a, b]$. If g is integrable on $[a, b]$, then prove that f is so and in this case

$$\int_a^b f\, dx = \int_a^b g\, dx$$

[*Hint*: $F(x) = f(x) - g(x) = 0$ except at a finite number of points of $[a, b]$, so that $F(x)$ has a finite number of points of discontinuity on $[a, b]$].

8. If $f(y, x) = 1 + 2x$, for y rational and $f(y, x) = 0$, for y irrational, find $F(y)$, where

$$F(y) = \int_0^1 f(y, x)\, dx$$

Is F integrable on $[0, 1]$?

9. Evaluate

$$\overline{\int_0^2} f(x)\, dx$$

$$\text{where } f(x) = \begin{cases} 0, & \text{when } x = n/(n+1), (n+1)/n, (n = 1, 2, 3 \ldots) \\ 1, & \text{elsewhere.} \end{cases}$$

Is f integrable on $[0, 2]$? Examine for continuity the function f so defined at the point $x = 1$.

10. f is bounded and integrable on $[a, b]$, show that

$$\int_a^b [f(x)]^2\, dx = 0$$

if and only if $f(c) = 0$, at every point, c, of continuity of f.

11. If f is integrable over $[a, b]$, under what conditions is $1/f$ integrable over $[a, b]$? State and prove a theorem about the integrability of $1/f$ over $[a, b]$.

12. A function f is defined on $[0, 1]$ as follows:

$$f(x) = \frac{1}{a^{r-1}} \text{ and } f(0) = 0, \text{ where } \frac{1}{a^r} < x \leq \frac{1}{a^{r-1}}, \text{ for } r = 1, 2, 3, \ldots,$$

where a is an integer greater than 2.

Show that $\int_0^1 f\, dx$ exists and is equal to $\dfrac{a}{a+1}$.

13. A function f is defined on $[0, 1]$, for positive integral value of r such that

$$f(x) = (-1)^{r-1}; \text{ where } 1/(r+1) < x < 1/r, (r = 1, 2, 3, \ldots)$$
$$f(0) = 0.$$

Show that $\int_0^1 f\, dx = \log 4 - 1$.

14. A function f is defined on [0, 1] by $f(x) = 2rx$, where $1/(r+1) < x < 1/r$, $(r = 1, 2, 3, \ldots)$, then show that $f \in R\ [0, 1]$ and its integral is $\pi^2/6$.

11. INTEGRATION BY PARTS

Theorem 20. *If f and g are integrable on $[a, b]$, and*

$$F(x) = A + \int_a^x f(x)\, dx,\ G(x) = B + \int_a^x g(x)\, dx$$

where A and B are constants, then

$$\int_a^b F(x)\, g(x)\, dx = \left[F(x)\, G(x)\right]_a^b - \int_a^b G(x)\, f(x)\, dx$$

[Here $\left[F(x)\, G(x)\right]_a^b$ denotes the difference $F(b)\, G(b) - F(a)\, G(a)$.]

Let $P = \{a = x_0, x_1, x_2, \ldots, x_n = b\}$ be a partition of $[a, b]$.

We have

$$\begin{aligned}\left[F(x)G(x)\right]_a^b &= \sum_{i=1}^{n} [F(x_i)\, G(x_i) - F(x_{i-1})\, G(x_{i-1})] \\ &= \sum G(x_i)[F(x_i) - F(x_{i-1})] + \sum F(x_{i-1})[G(x_i) - G(x_{i-1})] \\ &= \sum G(x_i) \int_{x_{i-1}}^{x_i} f(x)\, dx + \sum F(x_{i-1}) \int_{x_{i-1}}^{x_i} g(x)\, dx \qquad \ldots(1)\end{aligned}$$

Let $\Delta f_i = f(x_i) - f(x_{i-1})$, and $\Delta g_i = g(x_i) - g(x_{i-1})$ be the oscillation of f and g in Δx_i.

Now for all $x \in \Delta x_i$, we have

$$\begin{cases} \left| f(x) - f(x_i) \right| \le \left| f(x_i) - f(x_{i-1}) \right| = \Delta f_i \\ \left| g(x) - g(x_{i-1}) \right| \le \Delta g_i \end{cases}$$

$$\Rightarrow \quad \begin{cases} f(x_i) - \Delta f_i \le f(x) \le f(x_i) + \Delta f_i \\ g(x_{i-1}) - \Delta g_i \le g(x) \le g(x_{i-1}) + \Delta g_i \end{cases}$$

$$\Rightarrow \quad \begin{cases} [f(x_i) - \Delta f_i]\, \Delta x_i \le \displaystyle\int_{x_{i-1}}^{x_i} f(x)\, dx \le [f(x_i) + \Delta f_i]\, \Delta x_i \\ [g(x_{i-1}) - \Delta g_i]\, \Delta x_i \le \displaystyle\int_{x_{i-1}}^{x_i} g(x)\, dx \le [g(x_{i-1}) + \Delta g_i]\, \Delta x_i \end{cases}$$

CHAPTER 9

$$\Rightarrow \quad \begin{cases} \int_{x_i-1}^{x_i} f(x)\,dx = [f(x_i) + \theta_i\,\Delta f_i]\,\Delta x & ...(2) \\ \int_{x_i-1}^{x_i} g(x)\,dx = [g(x_i) + \theta_i\,\Delta g_i]\,\Delta x_i & ...(3) \end{cases}$$

where $-1 \le \theta_i, \theta_i' \le 1$.

Hence, from equations (1), (2) and (3), we get

$$[F(x)\,G(x)]_a^b = \Sigma\, G(x_i)\, f(x_i)\, \Delta x_i + \Sigma\, F(x_{i-1})\, g(x_{i-1})\, \Delta x_i + \sigma \qquad ...(4)$$

where

$$\sigma = \Sigma\, [G(x_i)\, \Delta f_i\, \theta_i + F(x_{i-1}) \Delta g_i\, \theta_i']\, \Delta x_i$$

Since F and G, being continuous, are bounded, therefore a number k exists such that

$$|F(x)| \le k, |G(x)| \le k, \quad \forall\, x \in [a, b]$$

$$\therefore \qquad |\sigma| \le k(\Sigma\, \Delta f_i + \Sigma\, \Delta g_i)\, \Delta x_i$$

In the limit when $\mu(P) \to 0$, $\sigma \to 0$, and (4) gives

$$[F(x)\,G(x)]_a^b = \int_a^b G(x)\, f(x)\; dx + \int_a^b F(x)\, g(x)\; dx$$

or
$$\int_a^b F(x)\, g(x)\; dx = [F(x)\,G(x)]_a^b - \int_a^b G(x)\, f(x)\; dx$$

Hence, the proof.

Corollary. If a function g is bounded and integrable on $[a, b]$ and if a function f is derivable and its derivative f' is bounded and integrable on $[a, b]$, then

$$\int_a^b f(x)\, g(x)\; dx = \left[f(x) \int_a^x g(x)\; dx \right]_a^b - \int_a^b \left\{ f'(x) \int_a^x g(x)\; dx \right\} dx$$

$$= f(b) \int_a^b g(x)\, dx - \int_a^b \left\{ f'(x) \int_a^x g(x)\; dx \right\} dx$$

If, however, both the derivatives f' and g' are assumed to be bounded and integrable, a much shorter and simpler proof exists, which follows.

Theorem 21. ***A particular case.*** *If f and g are both differentiable on $[a, b]$ and if f' and g' are both integrable on $[a, b]$ then*

$$\int_a^b f(x)\, g'(x)\; dx = [f(x) \cdot g(x)]_a^b - \int_a^b g(x)\, f'(x)\; dx$$

Since f and g are differentiable and hence continuous on $[a, b]$, therefore $f, g \in R[a, b]$. Again, since f, g, f' and g' are all integrable over $[a, b]$, therefore $fg', gf' \in R[a, b]$.

Let $F(x) = f(x)\, g(x)$, for all $x \in [a, b]$

$$\therefore \qquad F'(x) = f(x)g'(x) + g(x)\, f'(x)$$

$$\therefore \qquad \int_a^b F'(x)\, dx = \int_a^b \{f(x)g'(x) + g(x)\, f'(x)\}\, dx$$

$$= \int_a^b f(x)g'(x)\, dx + \int_a^b g(x)\, f'(x)\, dx \qquad \text{...(1)}$$

Also by Fundamental Theorem (17),

$$\int_a^b F'(x)\, dx = F(b) - F(a) = \left[f(x)g(x)\right]_a^b \qquad \text{...(2)}$$

From equations (1) and (2), we get

$$\int_a^b f(x)g'(x)\, dx = \left[f(x)g(x)\right]_a^b - \int_a^b g(x)f'(x)dx.$$

12. CHANGE OF VARIABLE IN AN INTEGRAL

Theorem 22. *If (i) $f \in R[a, b]$, (ii) ϕ is a derivable, strictly monotonic function on $[\alpha, \beta]$, where $a = \phi(\alpha), b = \phi(\beta)$, and (iii) $g' \in R[\alpha, \beta]$, then*

$$\int_a^b f(x)\, dx = \int_\alpha^\beta f(\phi(y))\, \phi'(y)\, dy$$

$$\left[\text{Change of variable in } \int_a^b f(x)\, dx \text{ by putting } x = \phi(y).\right]$$

Let ϕ be strictly monotonic increasing on $[\alpha, \beta]$.

Since ϕ is strictly monotonic, it is invertible, *i.e.*,

$$x = \phi(y) \qquad \Rightarrow y = \phi^{-1}(x), \;\; \forall\, x \in [a, b]$$

so that

$$\alpha = \phi^{-1}(a) \quad \text{and} \quad \beta = \phi^{-1}(b)$$

Let $P = \{a = x_0, x_1, x_2, \ldots, x_n = b\}$ be any partition of $[a, b]$ and

$Q = \{\alpha = y_0, y_1, y_2, \ldots, y_n = \beta\}$, $y_i = \phi^{-1}(x_i)$ the corresponding partition of $[\alpha, \beta]$.

By Lagrange's Mean Value Theorem,

$$\Delta x_i = \phi(y_i) - \phi(y_{i-1}) = \phi'(\eta_i)\,\Delta y_i, \quad \eta_i \in \Delta y_i \qquad \text{...(1)}$$

Let

$$\xi_i = \phi(\eta_i), \text{ where } \xi_i \in \Delta x_i \qquad \text{...(2)}$$

Now,

$$S(P, f) = \sum_{i=1}^{n} f(\xi_i)\,\Delta x_i = \sum_{i=1}^{n} f(\phi(\eta_i))\,\phi'(\eta_i)\,\Delta y_i \qquad \text{...(3)}$$

Uniform continuity of ϕ implies that $\mu(Q) \to 0$ as $\mu(P) \to 0$

Also $f \in R$ implies $\lim\limits_{\mu(P)\to 0} S(P, F)$ exists, and therefore the limit of the R.H.S. of (3) as $\mu(Q) \to 0$

also exists and equals the integral $\int_{\alpha}^{\beta} f(\phi(y))\,\phi'(y)\,dy$.

Hence, letting $\mu(P) \to 0$ in (3), we get

$$\int_a^b f(x)\,dx = \int_{\alpha}^{\beta} f(\phi(y))\,\phi'(y)\,dy.$$

Remarks:

1. If $\phi'(y) \neq 0$ for any $y \in [\alpha, \beta]$, then ϕ is strictly monotonic in $[\alpha, \beta]$. Hence, the condition of strict monotonicity of ϕ in the statement of the theorem, may be replaced by
$$\phi'(y) \neq 0, \quad \forall\ y \in [\alpha, \beta]$$
2. The theorem holds even if $\phi'(y) = 0$ for a finite number of values of y. For, in that case, the interval $[\alpha, \beta]$ can be divided into a finite number of sub-intervals, in each of which ϕ is strictly monotonic. Repetition of the argument for each of the sub-intervals in turn and the addition of the results, gives the required result.

13. SECOND MEAN VALUE THEOREM

The theorem is due to the great German mathematician, Karl Weierstrass, and the proof depends upon the *Abel's Lemma* and the *Bonnett's Theorem*. The theorem in fact is a generalisation of the Bonnett's Theorem, which, in reality, is the second mean value theorem under slightly more restricted conditions.

Abel's Lemma. *If $\{b_n\}$ is a positive monotone decreasing sequence and k, K denote respectively the least and the greatest values of the sums $\sum_{r=m}^{p} u_r$, for $p = m, m+1, \ldots, n$, then*

$$b_m k \leq \sum_{r=m}^{n} b_r u_r \leq b_m K$$

Let $S_p = \sum_{r=m}^{p} u_r$, then

$$\sum_{r=m}^{n} b_r u_r = b_m u_m + b_{m+1}u_{m+1} + \ldots + b_n u_n$$

$$= b_m S_m + b_{m+1}(S_{m+1} - S_m) + \ldots + b_n(S_n - S_{n-1})$$
$$= (b_m - b_{m+1})\, S_m + (b_{m+1} - b_{m+2})\, S_{m+1} + \ldots,$$
$$+ (b_{n-1} - b_n)\, S_{n-1} + b_n S_n$$

All brackets on the right are non-negative.

$\therefore$ $$k(b_m - b_{m+1} + b_{m+1} - b_{m+2} + \ldots + b_{n-1} - b_n + b_n)$$

$$\le \sum_{r=m}^{n} b_r u_r \le K(b_m - b_{m+1} + \ldots - b_n + b_n)$$

$\Rightarrow$ $$b_m k \le \sum_{r=m}^{n} b_r u_r \le b_m K$$

In particular ($m = 1$), the lemma may be stated in the form:

If $b_1, b_2, \ldots, b_n$ is a positive monotone decreasing set and k, K denote respectively the least and the greatest values of the partial sums, $\sum_{r=1}^{P} u_r,\ 1 \le p \le n$ of the numbers, $u_1, u_2, \ldots, u_n$, then

$$b_1 k \le \sum_{r=1}^{n} b_r u_r \le b_1 K.$$

Theorem 23. *Second mean value theorem. If $\int_a^b f\,dx$ and $\int_a^b g\,dx$ both exist and f is monotone on [a, b], then there exists $\xi \in [a, b]$ such that*

$$\int_a^b fg\,dx = f(a)\int_a^{\xi} g\,dx + f(b)\int_{\xi}^{b} g\,dx$$

We first prove the *Bonnett's form* of the theorem, where in addition to the hypothesis of the theorem, the monotone function is *positive* and *monotone decreasing* on $[a, b]$.

If $\int_a^b \phi\,dx$ and $\int_a^b g\,dx$ both exist and ϕ is positive and monotone decreasing on [a, b], then there exists a point $\xi \in [a, b]$ such that

$$\int_a^b \phi g\,dx = \phi(a)\int_a^{\xi} g\,dx$$

Let $P = \{a = x_0, x_1, \ldots, x_n = b\}$ be a partition of $[a, b]$, and M_i, m_i the bounds of g in Δx_i.

Let $t_1 = a$ and t_i $(i \ne 1)$ be any point of Δx_i.

In the sub-interval Δx_i, we know

CHAPTER 9

$$m_i \, \Delta x_i \le \int_{x_{i-1}}^{x_i} g \, dx \le M_i \, \Delta x_i$$

and

$$m_i \, \Delta x_i \le g(t_i) \, \Delta x_i \le M_i \, \Delta x_i$$

Letting $i = 1, 2, 3, \ldots, p$, $(p \le n)$ and adding vertically, we get

$$\sum_{i=1}^{p} m_i \, \Delta x_i \le \int_a^{xp} g \, dx \le \sum_{i=1}^{p} M_i \, \Delta x_i$$

and

$$\sum_{i=1}^{p} m_i \, \Delta x_i \le \sum_{i=1}^{p} g(t_i) \, \Delta x_i \le \sum_{i=1}^{p} M_i \, \Delta x_i$$

which gives

$$\left| \int_a^{xp} g \, dx - \sum_{i=1}^{p} g(t_i) \, \Delta x_i \right| \le \sum_{i=1}^{p} (M_i - m_i) \, \Delta x_i \le \omega(P, g)$$

$$\Rightarrow \qquad \int_a^{xp} g \, dx - \omega(P, g) \le \sum_{i=1}^{p} g(t_i) \, \Delta x_i \le \int_a^{xp} g \, dx + \omega(P, g)$$

where $\omega(P, g)$ denotes the oscillatory sum, $U(P, g) - L(P, g)$.

Now, $\int_a^t g \, dx$, being a continuous function (§ 8 Th. 16), is bounded. Let A, B be its bounds, so that we have

$$B - \omega(P, g) \le \sum_{i=1}^{p} g(t_i) \, \Delta x_i \le A + \omega(P, g) \qquad \ldots(1)$$

Using the Abel's lemma, where

$$b_i = \phi(t_i), \qquad u_i = g(t_i) \, \Delta x_i$$
$$k = B - \omega(P, g), \qquad K = A + \omega(P, g)$$

we get

$$\phi(a)[B - \omega(P, g)] \le \sum_{i=1}^{n} \phi(t_i) g(t_i) \, \Delta x_i \le \phi(a)[A + \omega(P, g)]$$

Taking the limit when $\mu(P) \to 0$, we get

$$B\phi(a) \le \int_a^b \phi g \, dx \le A\phi(a)$$

$$\Rightarrow \qquad \int_a^b \phi g \, dx = \mu \phi(a) \qquad \ldots(2)$$

where μ is some number between B and A.

The function $\int_a^b g(x)\,dx$ being continuous, must assume, for some $\xi \in [a, b]$, the value μ which lies between its bounds. Thus, we get

$$\int_a^b \phi g\,dx = \phi(a) \int_a^\xi g\,dx \qquad ...(3)$$

Let us now prove the theorem proper, due to Weierstrass.

Let, first, f be monotone *decreasing*, so that the function ϕ where $\phi = f - f(b)$ is positive and monotone decreasing. Therefore, by what has been proved above, there exists a number ξ between a and b such that

$$\int_a^b g[f - f(b)]\,dx = [f(a) - f(b)] \int_a^\xi g\,dx$$

$$\int_a^b fg\,dx = f(a) \int_a^\xi g\,dx + f(b) \left[\int_a^b g\,dx - \int_a^\xi g\,dx \right]$$

$$= (fa) \int_a^\xi g\,dx + f(b) \int_\xi^b g\,dx \qquad ...(4)$$

Let now f be monotone *increasing*, so that $(-f)$ is monotone decreasing and therefore from (4),

$$\int_a^b (-f)g\,dx = -f(a) \int_a^\xi g\,dx - f(b) \int_\xi^b g\,dx$$

$$\Rightarrow \qquad \int_a^b fg\,dx = f(a) \int_a^\xi g\,dx + f(b) \int_\xi^b g\,dx$$

Hence, the theorem.

Note: It may be easily verified that the theorem holds for $a > b$ also.

13.1 Second Mean Value Theorem (*A particular case*)

Theorem 24. *If f is monotonic and f, f' and g are all continuous in $[a, b]$, then there exists $\xi \in [a, b]$ such that*

$$\int_a^b f(x)g(x)\,dx = f(a) \int_a^\xi g(x)\,dx + f(b) \int_\xi^b g(x)\,dx$$

Let $G(x) = \int_a^x g(t)\,dt$

Clearly $G(a) = 0$, and under the given conditions, $G(x)$ is differentiable and $G'(x) = g(x)$.

$$\therefore \quad \int_a^b f(x)g(x)\,dx = \int_a^b f(x)G'(x)\,dx$$

$$= \left[f(x)G(x)\right]_a^b - \int_a^b G(x)f'(x)\,dx$$

on integrating by parts.

Since G, being continuous, is integrable and f is monotone and continuous on $[a, b]$, therefore on using generalised First Mean Value Theorem, $\exists\ \xi \in [a, b]$ such that

$$\int_a^b f(x)g(x)\,dx = f(b)G(b) - G(\xi)\int_a^b f'(x)\,dx$$

$$= f(b)G(b) - G(\xi)\{f(b) - f(a)\}$$

$$= f(b)\{G(b) - G(\xi)\} + f(a)\,G(\xi)$$

$$= f(b)\int_a^b g(x)\,dx + f(a)\int_a^\xi g(x)\,dx$$

Example 15. If a function f is continuous on [0, 1], show that

$$\lim_{x\to\infty}\int_0^1 \frac{nf(x)}{1+n^2x^2}\,dx = \frac{\pi}{2}f(0)$$

■ Let us put

$$\int_0^1 \frac{nf(x)}{1+n^2x^2}\,dx = \int_0^{1/\sqrt{n}} \frac{nf(x)}{1+n^2x^2}\,dx + \int_{1/\sqrt{n}}^1 \frac{nf(x)}{1+n^2x^2}\,dx$$

By generalised first mean value theorem,

$$\int_0^{1/\sqrt{n}} \frac{nf(x)}{1+n^2x^2}\,dx = f(\xi)\int_0^{1/\sqrt{n}} \frac{n}{1+n^2x^2}\,dx, \text{ where } 0 \le \xi \le \frac{1}{\sqrt{n}}$$

$$= f(\xi)\tan^{-1}\sqrt{n} \to \frac{\pi}{2}f(0) \quad \text{as } n\to\infty$$

Again, since f is continuous on [0, 1], it is bounded and therefore there exists K such that

$|f(x)| \le K, \quad \forall\ x \in [0, 1]$

$$\left|\int_{1/\sqrt{n}}^1 \frac{nf(x)}{1+n^2x^2}\,dx\right| = K\left|\int_{1/\sqrt{n}}^1 \frac{n}{1+n^2x^2}\,dx\right|$$

$$= K \,|\tan^{-1} n - \tan^{-1} \sqrt{n}$$
$$\to 0 \text{ as } n \to \infty$$

Hence, the result.

Example 16. *Riemann-Lebesgue Lemma.* If a function f is bounded and integrable on $[a, b]$, show that

$$\lim_{n\to\infty} \int_a^b f(x) \cos nx \, dx = 0 \text{ and } \lim_{n\to\infty} \int_a^b f(x) \sin nx \, dx = 0.$$

■ Let $I_n = \int_a^b f(x) \cos nx \, dx.$

Let, further, $\varepsilon > 0$ be an arbitrary number.

Since f is bounded and integrable on $[a, b]$, there exists a partition $P = \{a = x_0, x_1, x_2, \ldots, x_p = b\}$ such that the oscillatory sum

$$U(P, f) - L(P, f) = \sum_{i=1}^{p} (M_i - m_i)\,\Delta x_i < \varepsilon/2$$

where M_i, m_i are bounds of f in Δx_i.

Now

$$I_n = \sum_{i=1}^{p} \int_{x_{i-1}}^{x_i} f(x) \cos nx \, dx$$
$$= \sum_{i=1}^{p} f(x_{i-1}) \int_{x_{i-1}}^{x_i} \cos nx \, dx + \sum_{i=1}^{p} \int_{x_{i-1}}^{x_i} \{f(x) - f(x_{i-1})\} \cos nx \, dx$$

$$\therefore \qquad |I_n| \le \sum_{i=1}^{p} |f(x_{i-1})| \left| \int_{x_{i-1}}^{x_i} \cos nx \, dx \right|$$
$$+ \sum_{i=1}^{p} \int_{x_{i-1}}^{x_i} \left| \{f(x) - f(x_{i-1})\} \cos nx \right| dx$$

But for all $x \in \Delta x_i$, we have

$$|f(x) - f(x_{i-1})| \le M_i - m_i \ (i = 1, 2, \ldots, p)$$

so that
$$|\{f(x) - f(x_{i-1})\} \cos nx| \le M_i - m_i$$

and
$$\sum_{i=1}^{p} \int_{x_{i-1}}^{x_i} \left| \{f(x) - f(x_{i-1})\} \cos nx \right| dx \le \sum_{i=1}^{p} (M_i - m_i)(x_i - x_{i-1})$$
$$< \frac{1}{2}\varepsilon$$

Also
$$\left|\int_{x_{i-1}}^{x_i} \cos nx\, dx\right| = \left|\frac{1}{n}\{\sin nx_i - \sin nx_{i-1}\}\right|$$

$$\leq \frac{1}{n}\left\{\left|\sin nx_i\right| + \left|\sin nx_{i-1}\right|\right\} \leq \frac{2}{n}$$

$\therefore$
$$|I_n| \leq \frac{2}{n}\sum_{i=1}^{P}\left|f(x_{i-1})\right| + \frac{1}{2}\varepsilon$$

For a fixed $P, f(x_{i-1})$ is a fixed quantity. Also there exists a positive number m such that for all $n \geq m$,

$$\frac{2}{n}\sum_{i=1}^{P}\left|f(x_{i-1})\right| < \frac{1}{2}\varepsilon$$

$\therefore$
$$|I_n| < \varepsilon, \quad \forall\, n \geq m$$

$\Rightarrow$
$$\lim_{n\to\infty}\int_a^b f(x)\cos nx\, dx = 0$$

It may similarly be shown that

$$\lim_{n\to\infty}\int_a^b f(x)\sin nx\, dx = 0.$$

Notes:

1. For another proof see Fourier series.
2. In particular, if $f(x)$ is bounded and integrable in $\left[0, \frac{1}{2}\pi\right]$, then

$$\lim_{n\to\infty}\int_0^{\pi/2} f(x)\sin nx\, dx = 0.$$

If we assign the value 0 at the origin to $\left(\frac{1}{x} - \frac{1}{\sin x}\right)$, it becomes continuous, bounded and integrable in $\left[0, \frac{1}{2}\pi\right]$, so that

$$\lim_{n\to\infty}\int_0^{\pi/2} f(x)\left(\frac{1}{x} - \frac{1}{\sin x}\right)\sin nx\, dx = 0$$

$\Rightarrow$
$$\lim_{n\to\infty}\int_0^{\pi/2} f(x)\frac{\sin nx}{x}\, dx = \lim_{n\to\infty}\int_0^{\pi/2} f(x)\frac{\sin nx}{\sin x}\, dx.$$

Example 17. Show that $\lim I_n$, where

$$I_n = \int_0^{\delta} \frac{\sin nx}{x} dx, n \in N \text{ exists and equals } \pi/2$$

- Since $\lim\limits_{x \to 0} \frac{\sin nx}{x} = n$, the integrand becomes continuous for every value of x, if we assign to it the value n at $x = 0$.

I. Let us first prove that the integral exists (is convergent).

Let $I_n = \int_0^{\delta} \frac{\sin nx}{x} dx$

Putting $nx = t$, we get

$$I_n = \int_0^{n\delta} \frac{\sin t}{t} dt$$

$$\therefore \qquad |I_{n+p} - I_n| = \left| \int_{n\delta}^{(n+p)\delta} \frac{\sin t}{t} dt \right|, p \geq 1$$

$$\leq \int_{n\delta}^{(n+p)\delta} \frac{|\sin t|}{t} dt$$

Since $|\sin t|$ keeps the same sign (positive) and $1/t$ is positive and monotonic decreasing in $[n\delta, (n + p)\delta]$, using Bonnett's form of the second mean value theorem, we get

$$|I_{n+p} - I_n| \leq \frac{1}{n\delta} \int_{n\delta}^{(n+p)\delta} |\sin t| dt$$

$$\leq \frac{2}{n\delta} \leq \varepsilon, \forall n > \frac{2}{\varepsilon\delta}$$

Hence by Cauchy's principle of convergence, $\{I_n\}$ converges.

II. It will now be shown that

$$\lim_{n \to \infty} I_n = \lim_{n \to \infty} \int_0^{\pi/2} \frac{\sin nx}{\sin x} dx$$

Let us write

$$\int_0^{\pi/2} \frac{\sin nx}{x} dx = \int_0^{\delta} \frac{\sin nx}{x} dx + \int_{\delta}^{\pi/2} \frac{\sin nx}{x} dx$$

Using the preceding example,

$$\int_{\delta}^{\pi/2} \frac{\sin nx}{x}\, dx \to 0 \text{ as } n \to \infty$$

$$\Rightarrow \qquad \lim_{n\to\infty} I_n = \lim_{n\to\infty} \int_0^d \frac{\sin nx}{x}\, dx = \lim_{n\to\infty} \int_0^{\pi/2} \frac{\sin nx}{x}\, dx$$

The function f, where

$$f(x) = \begin{cases} \dfrac{1}{x} - \dfrac{1}{\sin x}, & x \neq 0 \\ 0, & x = 0 \end{cases}$$

is continuous and therefore bounded and integrable in $[0, \pi/2]$, so that using the preceding example, we get

$$\lim_{n\to\infty} \int_0^{\pi/2} \left(\frac{1}{x} - \frac{1}{\sin x}\right) \sin nx\, dx = 0$$

$$\Rightarrow \qquad \lim_{n\to\infty} I_n = \lim_{n\to\infty} \int_0^{\pi/2} \frac{\sin nx}{x}\, dx = \lim_{n\to\infty} \int_0^{\pi/2} \frac{\sin nx}{\sin x}\, dx.$$

III. Let us now proceed to evaluate $\lim I_n$ and this we do by letting $n \to \infty$ through odd integral values. We know that

$$\frac{\sin(2n+1)}{\sin x} = 2\left\{\frac{1}{2} + \cos 2x + \cos 4x + \ldots + \cos 2nx\right\}$$

Integrating, we get

$$\int_0^{\pi/2} \frac{\sin(2n+1)}{\sin x}\, dx = \frac{\pi}{2},\ \forall n \in \mathbf{N}$$

$$\lim I_n = \lim \int_0^{\pi/2} \frac{\sin x}{x}\, dx = \lim \int_0^{\pi/2} \frac{\sin nx}{x}\, dx$$

$$= \lim \int_0^{\pi/2} \frac{\sin(2x+1)}{\sin x}\, dx = \frac{\pi}{2}$$

Note: It will be defined later, in improper integrals, that

$$\int_0^\infty f\,dx = \lim_{x\to\infty} \int_0^x f\,dx$$

$$\int_0^x \frac{\sin x}{x}\,dx = \frac{x}{2}$$

EXERCISE

1. If f is continuous and positive on $[a, b]$, then show that $\int_a^b f\,dx$ is also positive.
2. Can the number ξ of the first mean value theorem always belong to $]a, b[$? Find the function f on some closed interval which satisfies the conditions of this theorem but for which ξ must be an end point of the interval.
3. How far can the Lagrange's mean value theorem for derivatives be used to make the statement : If f is the derivative of some function on $[a, b]$, then there exists a number $\xi \in [a,b]$ such that

$$\int_a^b f\,dx = f(\xi)\,(b-a)$$

[*Hint:* In Lagrange's mean value theorem, $\xi \in]a, b[$.]

4. If f is continuous on $[a, b]$ and $F(x) = \int_a^b f(t)\,dt$, for all $x \in [a, b]$, apply (if possible) the Lagrange's mean value theorem for derivatives to F over $[a, b]$. State the resulting theorem for **F.** Also state the first mean value theorem for integrals to $\int_a^b f\,dx$, and compare the two statements.

Also try to relate Cauchy mean value theorem with the generalised first mean value theorem for integrals.

5. If f and g are integrable and

$$\int_a^b f\,dx = \int_a^b g\,dx$$

then show that

$$f(\xi) = g(\xi), \text{ for some } \xi \in [a, b].$$

6. In the second mean value theorem show that f must be monotonic, by proving that the theorem does not hold in $\left[-\frac{1}{2}\pi, \frac{1}{2}\pi\right]$, if $f(x) = \cos x$, $g(x) = x^2$.

$$\left[\text{Note that } \int_{-\pi/2}^{\pi/2} x^2 \cos x\,dx > 0.\right]$$

7. Show that the second mean value theorem does not hold good in $[-1, 1]$, for $f(x) = g(x) = x^2$. Also test the validity of the first (generalised) mean value theorem.

8. Verify the two mean value theorems in [– 1, 1] for the functions $f(x) = e^x$, $g(x) = x$.

9. If $f \in R[a,b]$ and $F(x) = \int_a^x f(t)\,dt$, for all $x \in [a, b]$, then show that F is of bounded variation on $[a, b]$.

10. If $f, g \in R[a,b]$, then prove the following:

(*i*) $\left|\int_a^b fg\right| \le \left(\int_a^b f^2\right)^{1/2} \left(\int_a^b g^2\right)^{1/2}$, and

(*ii*) $\left(\int_a^b (f+g)^2\right)^{1/2} \le \left(\int_a^b f^2\right)^{1/2} + \left(\int_a^b g^2\right)^{1/2}$.

10 The Riemann-Stieltjes Integral

Having discussed the Riemann theory of integration to the extent possible within the scope of the present discussion, we now pass on to a generalisation of the subject. As mentioned earlier many refinements and extensions of the theory exist but we shall study briefly—in fact very briefly—the extension due to Stieltjes, known as the theory of *Riemann-Stieltjes integration*. The most noteworthy of the extensions, the *Lebesgue theory* of integration will be however discussed later in chapter 19.

It may be stated once for all that, unless otherwise stated, all functions will be *real-valued* and *bounded* on the domain of definition. The *function* α *will always be monotonic increasing*.

1. DEFINITIONS AND EXISTENCE OF THE INTEGRAL

Let f and α be bounded function on $[a, b]$ and α be *monotonic increasing* on $[a, b]$, $b \geq a$.

Corresponding to any partition

$$P = \{a = x_0, x_1, ..., x_n = b\}, \text{ of } [a, b]$$

we write

$$\Delta\alpha_i = \alpha(x_i) - \alpha(x_{i-1}), \quad i = 1, 2, ..., n.$$

It is clear that $\Delta\alpha_i \geq 0$. As in § 1.1 Ch. 9, we define two sums,

$$U(P, f, \alpha) = \sum_{i=1}^{n} M_i \Delta\alpha_i$$

$$L(P, f, \alpha) = \sum_{i=1}^{n} m_i \Delta\alpha_i$$

where m_i, M_i, are the bounds (infimum and supremum respectively) of f in Δx_i, respectively, called the *Upper* and the *Lower* sums of f corresponding to the partition P.

If m, M are respectively the lower and the upper bounds of f on $[a, b]$, we have

$$m \leq m_i \leq M_i \leq M$$

$$\Rightarrow \qquad m\,\Delta\alpha_i \leq m_i\,\Delta\alpha_i \leq M_i\,\Delta\alpha_i \leq M\,\Delta\alpha_i, \; \Delta\alpha_i \geq 0$$

Putting $i = 1, 2, ..., n$ and adding all inequalities, we get

$$m\{\alpha(b) - \alpha(a)\} \leq L(P, f, \alpha) \leq U(P, f, \alpha) \leq M\{\alpha(b) - \alpha(a)\} \qquad ...(1)$$

As in Riemann integration, § 1.1, we define two integrals, which always exist by a similar reasoning,

$$\overline{\int_a^b} f\, d\alpha = \inf.\, U(P, f, \alpha)$$

$$\underline{\int_a^b} f\, d\alpha = \sup.\, L(P, f, \alpha) \qquad \text{...(2)}$$

the infimum and supremum being taken over all partitions of $[a, b]$. These are, respectively, called the *upper* and the *lower* integrals of f with respect to α.

These two integrals may or may not be equal. In cases these two integrals are equal, *i.e.*,

$$\overline{\int_a^b} f\, d\alpha = \underline{\int_a^b} f\, d\alpha,$$

we say that f is integrable with respect to α in the Riemann sense and write $f \in R_\alpha\,[a, b]$ or simply $R(\alpha)$. Their common value is denoted by

$$\int_a^b f\, d\alpha$$

or sometimes by

$$\int_a^b f(x)\, d\alpha(x)$$

and is called the *Riemann-Stieltjes integral* (or simply the *Stieltjes integral*) of f with respect to α, over $[a, b]$.

From equations (1) and (2), it follows that

$$m\{\alpha(b) - \alpha(a)\} \le L(P, f, \alpha) \le \underline{\int_a^b} f\, d\alpha \le \overline{\int_a^b} f\, d\alpha$$

$$\le U(P, f, \alpha) \le m\{\alpha(b) - \alpha(a)\} \qquad \text{...(3)}$$

Remark: The upper and the lower integrals always exist for bounded functions but these may not be equal for all bounded functions. Such functions are not integrable. Thus the question of their equality and hence that of the integrability of the function is our main concern.

The Riemann-Stieltjes integral reduces to Riemann integral when $\alpha(x) = x$.

1.1 Some Deductions

(*i*) If $f \in R(\alpha)$, then $\exists$ a number λ lying between the bounds of f such that

$$\int_a^b f\, d\alpha = \lambda\{a(b) - \alpha(a)\} \qquad \text{(using 3)}$$

(*ii*) If f is continuous on $[a, b]$, then $\exists$ a number $\xi \in [a, b]$ such that

$$\int_a^b f\, d\alpha = f(\xi)\,\{\alpha(b) - \alpha(a)\}$$

(*iii*) If $f \in R(\alpha)$, and k is a number such that

$$|f(x)| \le k, \text{ for all } x \in [a, b]$$

then

$$\left|\int_a^b f\, d\alpha\right| \le k\{\alpha(b) - \alpha(b)\}$$

(*iv*) If $f \in R(\alpha)$, over $[a, b]$ and $f(x) \ge 0$, for all $x \in [a, b]$, then

$$\int_a^b f\, d\alpha \begin{cases} \ge 0, & b \ge a \\ \le 0, & b \le a \end{cases}$$

Since $f(x) \ge 0$, the lower bound $m \ge 0$ and therefore the result follows from equations (3).

(*v*) If $f \in R(\alpha)$, $g \in R(\alpha)$ over $[a, b]$ such that $f(x) \ge g(x)$, then

$$\int_a^b f\, d\alpha \ge \int_a^b f\, d\alpha, b \ge a$$

and

$$\int_a^b f\, d\alpha \le \int_a^b g\, d\alpha, b \le a$$

The result follows by reasoning similar to that of Deduction 5 § 1.4, Chapter 9.

1.2 Refinement of Partitions

Theorem 1. *If P^* is a refinement of P, then*

(*i*) $L(P^*, f, \alpha) \ge L(P, f, \alpha)$, *and*

(*ii*) $L(P^*, f, \alpha) \le U(P, f, \alpha)$.

Let us prove (*ii*).

Let $P = \{a = x_0, x_1, ..., x_n = b\}$ be a partition of the given interval. Suppose first that P^* contains just one point more than P. Let this extra point ξ belongs to Δx_i, *i.e.*, $x_{i-1} < \xi < x_i$.

As f is bounded over the entire interval $[a, b]$, it is bounded on every sub-interval Δx_i $(i = 1, 2, ..., n\}$. Let W_1, W_2, M_i be the upper bounds (supremum) of f in the intervals $[x_{i-1}, \xi]$, $[\xi, x_i]$, $[x_{i-1}, x_i]$, respectively.

Clearly $\quad W_1 \le M_i, \; W_2 \le M_i$.

$$\begin{aligned}\therefore\; U(P^*, f, \alpha) - U(P, f, \alpha) &= W_1\{\alpha(\xi) - \alpha(x_{i-1})\} + W_2\{\alpha(x_i) - \alpha(\xi)\} - M_i\{\alpha(x_i) - \alpha(x_{i-1})\} \\ &= (W_1 - M_i)\{a(\xi) - \alpha(x_{i-1})\} + (W_2 - M_i)\{a(x_i) - \alpha(\xi)\} \le 0\end{aligned}$$

$$\Rightarrow \quad U(P^*, f, \alpha) \le U(P, f, \alpha)$$

If P^* contains m points more than P, we repeat the above reasoning m times and arrive at the result (*ii*). The proof of (*i*) is similar.

Theorem 2. *For any two partitions* P_1, P_2,

$$L(P_1, f, \alpha) \leq U(P_2, f, \alpha)$$

i.e., no upper sum can ever be less than any lower sum.

Corollary. For a *bounded function f,*

$$\underline{\int_a^b} f\, d\alpha x \leq \overline{\int_a^b} f\, d\alpha$$

The proofs are similar to that of Theorem 9.2, Chapter 9.

Ex. If $P^* \supseteq P$, then show that

$$U(P^*, f, \alpha) - L(P^*, f, \alpha) \leq U(P, f, \alpha) - L(P, f, \alpha)\,.$$

2. A CONDITION OF INTEGRABILITY

Theorem 3. *A function f is integrable with respect to* α *on* $[a, b]$ *if and only if for every* $\varepsilon > 0$ *there exists a partition P of* $[a, b]$ *such that*

$$U(P, f, \alpha) - L(P, f, \alpha) < \varepsilon$$

Necessary. Let $f \in R(\alpha)$ over $[a, b]$.

$$\therefore \qquad \underline{\int_a^b} f\, d\alpha = \overline{\int_a^b} f\, d\alpha = \int_a^b f\, d\alpha$$

Let $\varepsilon > 0$ be any number.

Since the upper and the lower integrals are the infimum and the supremum, respectively, of the upper and the lower sums, therefore $\exists$ partitions P_1 and P_2 such that

$$U(P_1, f, \alpha) < \overline{\int_a^b} f\, d\alpha + \frac{1}{2}\varepsilon = \int_a^b f\, d\alpha + \frac{1}{2}\varepsilon$$

$$L(P_2, f, \alpha) > \underline{\int_a^b} f\, d\alpha - \frac{1}{2}\varepsilon = \int_a^b f\, d\alpha - \frac{1}{2}\varepsilon$$

Let $P = P_1 \cup P_2$ be the common refinement of P_1 and P_2.

$$\therefore \qquad U(P, f, \alpha) \leq U(P_1, f, \alpha)$$

$$< \int_a^b f\, d\alpha + \frac{1}{2}\varepsilon < L(P_2, f, \alpha) + \varepsilon \leq L(P, f, \alpha) + \varepsilon$$

$$\Rightarrow \qquad U(P, f, \alpha) - L(P, f, \alpha) < \varepsilon$$

Sufficient. For $\varepsilon > 0$, let P be a partition for which

$$U(P, f, \alpha) - L(P, f, \alpha) < \varepsilon$$

For any partition P, we know that

$$L(P, f, \alpha) \le \underline{\int_a^b} f\, d\alpha \le \overline{\int_a^b} f\, d\alpha \le U(P, f, \alpha)$$

$$\therefore \qquad \overline{\int} f\, d\alpha - \underline{\int} f\, d\alpha \le U(P, f, \alpha) - L(P, f, \alpha) < \varepsilon$$

But a non-negative number can be less than every positive number, if it is zero.

$$\therefore \qquad \overline{\int_a^b} f\, d\alpha - \underline{\int_a^b} f\, d\alpha$$

so that $f \in R(\alpha)$, over $[a, b]$.

3. SOME THEOREMS

(*a*) *If* $f_1 \in R(\alpha)$ *and* $f_2 \in R(\alpha)$ over $[a, b]$, then

$$f_1 + f_2 \in R(\alpha) \text{ and } \int_a^b (f_1 + f_2)\, d\alpha = \int_a^b f_1\, d\alpha + \int_a^b f_2\, d\alpha$$

(*b*) *If* $f \in R(\alpha)$, *and c is a constant, then*

$$cf \in R(\alpha) \text{ and } \int_a^b cf\, d\alpha = c \int_a^b f\, d\alpha$$

(*c*) *If* $f_1 \in R(\alpha)$, $f_2 \in R(\alpha)$ *and* $f_1(x) \le f_2(x)$ *on* $[a, b]$, *then*

$$\int_a^b f_1\, d\alpha \le \int_a^b f_2\, d\alpha$$

(*d*) If $f \in R(\alpha)$ *over* $[a, b]$ *and if* $a < c < b$, *then*

$$f \in R(\alpha) \text{ on } [a, c], \text{ and on } [c, b] \text{ and } \int_a^c f\, d\alpha + \int_c^b f\, d\alpha = \int_a^b f\, d\alpha$$

(*e*) If $f \in R(\alpha)$ *over* $[a, b]$, *then*

$$|f| \in R(\alpha) \text{ and } \left| \int_a^b f\, d\alpha \right| \le \int_a^b |f|\, d\alpha$$

(*f*) *If* $f \in R(\alpha)$ *on* $[a, b]$, *then*

$$f^2 \in R(\alpha)$$

(g) *If* $f \in R(\alpha_1)$ *and* $f \in R(\alpha_2)$, *then*

$$f \in R(\alpha_1 + \alpha_2) \text{ and } \int_a^b f \, d(\alpha_1 + \alpha_2) = \int_a^b f \, d\alpha_1 + \int_a^b f \, d\alpha_2$$

and if $f \in R(\alpha)$ *and c a positive constant, then*

$$f \in R(c\alpha) \text{ and} \int_a^b f \, d(c\alpha) = c \int_a^b f \, d\alpha.$$

(*a*) Let $f = f_1 + f_2$.

Clearly f is bounded on $[a, b]$.

If $P = \{a = x_0, x_1 \ldots, x_n = b\}$ be any partition of $[a, b]$ and $m_i', M_i'; m_i'', M_i''; m_i, M_i$ the bounds of f_1, f_2 and f, respectively, on Δx_i, then

$$m_i' + m_i'' \le m_i \le M_i \le M_i' + M_i''$$

Multiplying by $\Delta\alpha_i$ and adding all these inequalities for $i = 1, 2, 3, \ldots, n$, we get

$$L(P, f_1, \alpha) + L(P, f_2, \alpha) \le (P, f, \alpha) \le U(P, f, \alpha)$$
$$\le U(P, f_1, \alpha) + U(P, f_2, \alpha) \qquad \ldots(1)$$

Let $\varepsilon > 0$ be any number.

Since $f_1 \in R(\alpha), f_2 \in R(\alpha)$, therefore $\exists$ partitions P_1, P_2 such that

$$U(P_1, f_1, \alpha) - L(P_1, f_1, \alpha) < \frac{1}{2}\varepsilon$$

$$U(P_2, f_2, \alpha) - L(P_2, f_2, \alpha) < \frac{1}{2}\varepsilon$$

Let $P = P_1 \cup P_2$, a refinement of P_1 and P_2.

$\therefore$
$$U(P, f_1, \alpha) - L(P, f_1, \alpha) < \frac{1}{2}\varepsilon$$

$$U(P, f_2, \alpha) - L(P, f_2, \alpha) < \frac{1}{2}\varepsilon \qquad \ldots(2)$$

Thus for partition P, we get from (1) and (2),

$$U(P, f, \alpha) - L(P, f, \alpha) \le U(P, f_1, \alpha) + U(P, f_2, \alpha) - L(P, f_1, \alpha) - L(P, f_2, \alpha)$$
$$< \frac{1}{2}\varepsilon + \frac{1}{2}\varepsilon = \varepsilon$$

$\Rightarrow$ $f \in R(\alpha)$ over $[a, b]$

Let us now proceed to prove the second part.

Since the upper integral is the infimum of the upper sums, therefore $\exists$ partitions P_1, P_2 such that

$$U(P_1, f_1, \alpha) < \int_a^b f_1 \, d\alpha + \frac{1}{2}\varepsilon$$

$$U(P_2, f_2, \alpha) < \int_a^b f_2 \, d\alpha + \frac{1}{2}\varepsilon$$

If $P = P_1 \cup P_2$, we have

$$\left.\begin{aligned} U(P, f_1, \alpha) &< \int_a^b f_1\, d\alpha + \frac{1}{2}\varepsilon \\ U(P, f_2, \alpha) &< \int_a^b f_2\, d\alpha + \frac{1}{2}\varepsilon \end{aligned}\right\} \quad \text{...(3)}$$

For such a partition P,

$$\int_a^b f\, d\alpha \le U(P, f, \alpha) \le U(P, f_1, \alpha) + U(P, f_2, \alpha) \quad \text{[from (1)]}$$

$$\le \int_a^b f_1\, d\alpha + \int_a^b f_2\, d\alpha + \varepsilon \quad \text{[using (3)]}$$

Since ε is arbitrary, we get

$$\int_a^b f\, d\alpha \le \int_a^b f_1\, d\alpha + \int_a^b f_2\, d\alpha \quad \text{...(4)}$$

Proceeding with $(-f_1)$ and $(-f_2)$ instead of f_1 and f_2, we get

$$\int_a^b f\, d\alpha \ge \int_a^b f_1\, d\alpha + \int_a^b f_2\, d\alpha \quad \text{...(5)}$$

(4) and (5) give

$$\int_a^b f\, d\alpha = \int_a^b f_1\, d\alpha + \int_a^b f_2\, d\alpha$$

(*g*) Since $f \in R(\alpha_1)$ and $f \in R(\alpha_2)$, therefore for $\varepsilon > 0$, $\exists$ partitions P_1, P_2 of $[a, b]$ such that

$$U(P_1, f, \alpha_1) - L(P_1, f, \alpha_1) < \frac{1}{2}\varepsilon$$

$$U(P_2, f, \alpha_2) - L(P_2, f, \alpha_2) < \frac{1}{2}\varepsilon$$

Let $P = P_1 \cup P_2$

$\therefore$

$$U(P, f, \alpha_1) - L(P, f, \alpha_1) < \frac{1}{2}\varepsilon \quad \text{...(1)}$$

$$U(P, f, \alpha_2) - L(P, f, \alpha_2) < \frac{1}{2}\varepsilon$$

Let the partition P be $\{a = x_0, x_1, x_2, ..., x_n = b\}$, and m_i, M_i be bounds of f in Δx_i.

Let $\alpha = \alpha_1 + \alpha_2$.

$\therefore$

$$\alpha(x) = \alpha_1(x) + \alpha_2(x)$$

$$\Delta x_{1i} = \alpha_1(x_i) - \alpha_1(x_{i-1})$$

$$\Delta x_{2i} = \alpha_2(x_i) - \alpha_2(x_{i-1})$$

$$\therefore \qquad \Delta x_i = \alpha(x_i) - \alpha(x_{i-1})$$

$$= \alpha_1(x_i) + \alpha_2(x_i) - \alpha_1(x_{i-1}) - \alpha_2(x_{i-1})$$

$$= \Delta\alpha_{1i} + \Delta\alpha_{2i}$$

$$\therefore \qquad U(P, f, \alpha) = \sum_i M_i \,\Delta\alpha_i$$

$$= \sum_i M_i \,(\Delta\alpha_{1i} + \Delta\alpha_{2i})$$

$$= U(P, f, \alpha_1) + U(P, f, \alpha_2) \qquad \text{...(2)}$$

Similarly,

$$L(P, f, \alpha) = L(P, f, \alpha_1) + L(P, f, \alpha_2) \qquad \text{...(3)}$$

$$\therefore \qquad U(P, f, \alpha) - L(P, f, \alpha) = U(P, f, \alpha_1) - L(P, f, \alpha_1)$$

$$+ U(P, f, \alpha_2) - L(P, f, \alpha_2)$$

$$< \frac{1}{2}\varepsilon + \frac{1}{2}\varepsilon = \varepsilon \qquad \text{[using (1)]}$$

$$\Rightarrow \qquad f \in R(\alpha), \text{ where } \alpha = \alpha_1 + \alpha_2$$

Now to prove the second part, we notice that

$$\int_a^b f \, d\alpha = \inf U(P, f, \alpha)$$

$$= \inf \{U(P, f, \alpha_1) + U(P, f, \alpha_2)\}$$

$$\geq \inf U(P, f, \alpha_1) + \inf U(P, f, \alpha_2)$$

$$= \int_a^b f \, d\alpha_1 + \int_a^b f \, d\alpha_2 \qquad \text{...(4)}$$

Similarly,

$$\int_a^b f \, d\alpha = \sup L(P, f, \alpha)$$

$$\leq \int_a^b f \, d\alpha_1 + \int_a^b f \, d\alpha_2 \qquad \text{...(5)}$$

From equations (4) and (5),

$$\int_a^b f \, d\alpha = \int_a^b f \, d\alpha_1 + \int_a^b f \, d\alpha_2$$

where $\alpha = \alpha_1 + \alpha_2$.

The proofs of the remaining parts are so similar to the above proofs and virtually identical to those of the corresponding theorems for Riemann integral that it is a mere repetition and are therefore left to the reader.

Corollary. If $f_1 \in R(\alpha)$ and $f_2 \in R(\alpha)$ over $[a, b]$, then

$$f_1 \cdot f_2 \in R(\alpha)$$

We know that if f_1, f_2 are integrable then $f_1 + f_2$, $f_1 - f_2$, f_1^2, f_2^2, are all integrable.

Also, then $(f_1 + f_2)^2$, $(f_1 - f_2)^2$ are integrable.

Now

$$4f_1 \cdot f_2 = (f_1 + f_2)^2 - (f_2 - f_2)^2$$

$$\Rightarrow \qquad f_1 \cdot f_2 \in R(\alpha)$$

4. A DEFINITION (*Integral as a limit of sum*)

As an analog to the Riemann sum, we introduce a sum which will lead to a sufficient condition for the existence of a Riemann-Stieltjes integral.

Definition. Corresponding to a partition P of $[a, b]$ and $t_i \in \Delta x_i$, consider the sum

$$S(P, f, \alpha) = \sum_{i=1}^{n} f(t_i)\, \Delta\alpha_i$$

We say that $S(P, f, \alpha)$ converges to A as $\mu(P) \to 0$, *i.e.*,

$$\lim_{\mu(P)\to 0} S(P, f, \alpha) = A$$

if, for every $\varepsilon > 0$ there exists $\delta > 0$ such that $|S(P, f, \alpha) - A| < \varepsilon$, for every partition $P = \{a = x_0, x_1, x_2, ..., x_n = b\}$, of $[a, b]$, with mesh $\mu(P) < \delta$ and every choice of t_i in Δx_i.

Theorem 4. *If* $\lim S(P, f, \alpha)$ *exists as* $\mu(P) \to 0$, *then*

$$f \in R(\alpha), \text{ and } \lim_{\mu(p)\to 0} S(P, f, \alpha) = \int_a^b f\, d\alpha$$

Let us suppose that $\lim S(P, f, \alpha)$ exists as $\mu(P) \to 0$ and is equal to A.

Therefore for $\varepsilon > 0, \exists\, \delta > 0$ such that for every partition P of $[a, b]$ with mesh $\mu(P) \to \delta$ and every choice of t_i and Δx_i, we have

$$|S(P, f, \alpha) - A| < \frac{1}{2}\varepsilon$$

or

$$A - \frac{1}{2}\varepsilon < S(P, f, \alpha) < A + \frac{1}{2}\varepsilon \qquad ...(1)$$

Let P be one such partition. If we let the points t_i range over the intervals Δx_i and take the infimum and the supremum of the sums $S(P, f, \alpha)$, (1) yields

$$A - \frac{1}{2}\varepsilon < L(P, f, \alpha) \le U(P, f, \alpha) < A + \frac{1}{2}\varepsilon \quad \text{...(2)}$$

$$\Rightarrow \quad U(P, f, \alpha) - L(P, f, \alpha) < \varepsilon$$

$$\Rightarrow \quad f \in R(\alpha) \text{ over } [a, b]$$

Again, since $S(P, f, \alpha)$ and $\int_a^b f\,d\alpha$ lie between $U(P, f, \alpha)$ and $L(P, f, \alpha)$

$$\therefore \quad \left| S(P, f, \alpha) - \int_a^b f\,d\alpha \right| \le U(P, f, \alpha) - L(P, f, \alpha) < \varepsilon$$

$$\Rightarrow \quad \lim_{\mu(p)\to 0} S(P, f, \alpha) = \int_a^b f\,d\alpha$$

Remark: The theorem asserts that the existence of the limit of $S(P, f, \alpha)$ implies that $f \in R(\alpha)$. The existence of the limit is a sufficient condition for $f \in R(\alpha)$ but as shown in Example 3 it is not a necessary condition, *i.e.*, functions exist which are integrable but for which limit of $S(P, f, \alpha)$ does not exist. Thus, whenever $\lim S(P, f, \alpha)$ exists, it is equal to $\int f\,d\alpha$. But when $f \in R(\alpha)$ nothing can be said about the existence of $\lim S(P, f, \alpha)$.

Theorem 5. *If f is continuous on $[a, b]$ then $f \in R(\alpha)$ over $[a, b]$. Moreover, to every $\varepsilon > 0$ there corresponds a $\delta > 0$ such that*

$$\left| S(P, f, \alpha) - \int_a^b f\,d\alpha \right| < \varepsilon$$

for every partition $P = \{a = x_0, x_1, x_2, ..., x_n = b\}$ of $[a, b]$ with $\mu(P) < \delta$, and for every choice of t_i in Δx_i, i.e.,

$$\lim_{\mu(p)\to 0} S(P, f, \alpha) = \int_a^b f\,d\alpha$$

[We still assume that all functions are bounded and α is monotonic increasing.]

Let $\varepsilon > 0$ be given, and let us choose $\eta > 0$ such that

$$\eta\{\alpha(b) - \alpha(a)\} < \varepsilon \quad \text{...(1)}$$

Since continuity of f on the closed interval $[a, b]$ implies its uniform continuity on $[a, b]$, therefore for $\eta > 0$ there corresponds $\delta > 0$ such that

$$|f(t_1) - f(t_2)| < \eta, \text{ if } |t_1 - t_2| < \delta, t_1, t_2 \in [a, b] \quad \text{...(2)}$$

Let P be a partition of $[a, b]$, with norm $\mu(P) < \delta$.

Then in view of (2),

$$M_i - m_i \le \eta,\ i = 1, 2, ..., n$$

$$\therefore \quad U(P, f, \alpha) - L(P, f, \alpha) = \sum_i (M_i - m_i)\, \Delta x_i$$

$$\le \eta \sum_i \Delta x_i$$

$$= \eta\{\alpha(b) - \alpha(a)\} < \varepsilon \qquad ...(3)$$

$$\Rightarrow \quad f \in R(\alpha) \text{ over } [a, b].$$

Again if $f \in R(\alpha)$, then for $\varepsilon > 0, \exists\, \delta > 0$ such that for all partitions P with $\mu(P) < \delta$,

$$\left|U(P, f, \alpha) - L(P, f, \alpha)\right| < \varepsilon$$

Since $S(P, f, \alpha)$ and $\int_a^b f\, d\alpha$ both lie between $U(P, f, \alpha)$ and $L(P, f, \alpha)$ for all partitions P with $\mu(P) < \delta$ and for all positions of t_i *in* Δx_i.

$$\therefore \quad \left|S(P, f, \alpha) - \int_a^b f\, d\alpha\right| < U(P, f, \alpha) - L(P, f, \alpha) < \varepsilon$$

$$\Rightarrow \quad \lim_{\mu(p)\to 0} S(P, f, \alpha) = \lim_{\mu(p)\to 0} \sum_{i=1}^{n} f(t_i)\, \Delta\alpha_i = \int_a^b f\, d\alpha$$

Notes:

1. Continuity is a sufficient condition for integrability of a function. It is not necessary condition. Functions exist which are integrable but not continuous.
2. For continuous function f, lim $S(P, f, \alpha)$ exists and equals $\int f\, d\alpha$.

Theorem 6. *If f is monotonic on* $[a, b]$, *and if* α *is continuous on* $[a, b]$, *then* $f \in R(\alpha)$.

[Monotonicity of α is still assumed.]

Let $\varepsilon > 0$ be a given positive number.

For any positive integer n, choose a partition $P = \{x_0, x_1, ..., x_n\}$ of $[a, b]$ such that

$$\Delta\alpha_i = \frac{\alpha(b) - \alpha(a)}{n},\ i = 1, 2, ..., n$$

This is possible because α is continuous and monotonic increasing on the closed interval $[a, b]$ and thus assumes every value between its bounds, $\alpha(a)$ and $\alpha(b)$.

Let f be monotonic increasing on $[a, b]$, so that its lower and the upper bound, m_i, M_i in Δx_i are given by

$$m_i = f(x_{i-1}),\ M_i = f(x_i);\ i = 1, 2, ..., n$$

$$\therefore \quad U(P, f, \alpha) - L(P, f, \alpha) = \sum_{i=1}^{n} (M_i - m_i)\, \Delta x_i$$

$$= \frac{\alpha(b) - \alpha(a)}{n} \sum_{i=1}^{n} \{f(x_i) - f(x_{i-1})\}$$

$$= \frac{\alpha(b) - \alpha(a)}{n} \{f(b) - f(a)\}$$

$$< \varepsilon, \text{ for large } n$$

$$\Rightarrow \quad f \in R(\alpha) \text{ over } [a, b]$$

Note: $f \in R(\alpha)$, *i.e.*, $\int f\, d\alpha$ exists when either

(*i*) f is continuous and α is monotonic, or

(*ii*) f is monotonic and α is continuous; of course α is still monotonic.

4.1 Some Examples

Example 1. A function α increases on $[a, b]$ and is continuous at x' where $a \le x' \le b$. Another function f is such that

$$f(x') = 1, \text{ and } f(x) = 0, \text{ for } x \ne x'$$

Prove that

$$f \in R(\alpha) \text{ over } [a, b], \text{ and } \int_a^b f\, d\alpha = 0.$$

■ Let $P = \{a = x_0, x_1, x_2, ..., x_n = b\}$ be a partition of $[a, b]$ and let $x' \in \Delta x_i$.

But since α is continuous at x' and increases on $[a, b]$, therefore for $\varepsilon > 0$ we can choose $\delta > 0$ such that

$$\Delta\alpha_i = \alpha(x_i) - \alpha(x_{i-1}) < \varepsilon, \text{ for } \Delta x_i < \delta$$

Let P be a partition with $\mu(P) < \delta$. Now

$$U(P, f, \alpha) = \Delta\alpha_i$$
$$L(P, f, \alpha) = 0$$

$$\therefore \quad \overline{\int_a^b} f\, d\alpha = \inf U(P, f, \alpha), \text{ over all partitions } P \text{ with } \mu(P) < \delta$$

$$= 0 = \underline{\int_a^b} f\, d\alpha$$

$$\Rightarrow \quad f \in R(\alpha), \text{ and } \int_a^b f\, d\alpha = 0.$$

Aliter. Let $P = \{a = x_0, x_1, ..., x_n = b\}$ be a partition of $[a, b]$ and let $x' \in \Delta x_i$, $x_{i-1} \le x' > x_i$.

By continuity of α at x', for $\varepsilon > 0$, $\exists\, \delta > 0$ such that

$$|\alpha(x) - \alpha(x')| < \frac{1}{2}\varepsilon, \text{ for } |x - x'| < \delta$$

Again, since α is an increasing function, we have

$$\alpha(x) - \alpha(x') < \frac{1}{2}\varepsilon, \text{ for } 0 < x - x' < \delta$$

and $$\alpha(x') - \alpha(x) < \frac{1}{2}\varepsilon, \text{ for } 0 < x' - x < \delta$$

Let P be a partition with $\mu(P) < \delta$.

$$\therefore \quad \begin{aligned} \Delta\alpha_i &= \alpha(x_i) - \alpha(x_{i-1}) \\ &= \alpha(x_i) - \alpha(x') + \alpha(x') - \alpha(x_{i-1}) \\ &< \frac{1}{2}\varepsilon + \frac{1}{2}\varepsilon = \varepsilon \end{aligned}$$

$$S(P, f, \alpha) = \sum_{i=1}^{n} f(t_i)\,\Delta\alpha_i = f(t_i)\,\Delta\alpha_i$$

$$= \begin{cases} 0, & t_i \ne x' \\ \Delta\alpha_i, & t_i = x' \end{cases}$$

$$\therefore \quad |S(P, f, \alpha)| = 0, \text{ when } t_i \ne x'$$
$$< \varepsilon, \text{ when } t_i = x'$$

In either case

$$\lim_{\mu(P)\to 0} S(P, f, \alpha) = 0$$

$$\Rightarrow \quad f \in R(\alpha) \text{ over } [a, b], \text{ and } \int_a^b f\,d\alpha = 0.$$

Example 2. f is a function bounded on $[-1, 1]$, and three functions $\beta_1, \beta_2, \beta_3$ are defined as follows:

$$\beta_1(x) = \begin{cases} 0, & x \le 0 \\ 1, & x > 0 \end{cases}$$

$$\beta_2(x) = \begin{cases} 0, & x < 0 \\ 1, & x \ge 0 \end{cases}$$

$$\beta_3(x) = \begin{cases} 0, & x < 0 \\ \frac{1}{2}, & x = 0 \\ 1, & x > 0 \end{cases}$$

Prove that $f \in R(\beta_3)$ iff f is continuous at $x = 0$, and then

$$\int_{-1}^{1} f\,d\beta_3 = f(0).$$

■ Let $P = \{-1 = x_0, x_1, ..., x_{i-2}, 0 = x_{i-1}, x_i, ..., x_n = 1\}$ be a partition of $[-1, 1]$ such that $x_{i-1} = 0$. Let $t_i \in \Delta x_i$.

Now

$$S(P, f, \beta_3) = \sum_{j=1}^{n} f(t_j)\{\beta_3(x_j) - \beta_3(x_{j-1})\}$$

$$= f(t_{i-1}) \cdot \frac{1}{2} + f(t_i) \cdot (1 - \frac{1}{2})$$

$$= \frac{1}{2}\{f(t_{i-1}) + f(t_i)\} \quad ...(1)$$

$$= f(0) \text{ in particular when } t_{i-1} = 0 = t_i \quad ...(2)$$

Clearly t_{i-1} tends to 0 from below and t_i from above, when the norm $\mu(P)$ tends to zero.

Hence $\lim_{\mu(P)\to 0} S(P, f, \alpha)$ exists when both the limits, $\lim_{t_{i-1}\to -0} f(t_{i-1})$ and $\lim_{t_i \to 0} f(t_i)$ or equivalently $\lim_{x\to 0-0} f(x)$ and $\lim_{x\to 0+0} f(x)$ exist, *i.e.*, both $f(0-)$ and $f(0+)$ exist.

Moreover, from (2) it is evident that these limits are each equal to $f(0)$. In that case

$$\lim_{\mu(P)\to 0} S(P, f, \alpha) = f(0)$$

Hence $f \in R(\beta_3)$ if $f(0+) = f(0-) = f(0)$, *i.e.*, if the function f is continuous at zero and in that case

$$\int_{-1}^{1} f \, d\beta_3 = f(0)$$

Also it is clear that f is continuous if $\lim S(P, f, \alpha)$ exists. Hence $f \in R(\beta_3)$ iff f is continuous at $x = 0$.

Example 3. For the functions β_1 and β_2 defined in Example 2, prove that $\beta_2 \in R(\beta_1)$, although $\lim S(P, \beta_2, \beta_1)$ does not exist, as $\mu(P) \to 0$.

■ Let $P = \{-1 = x_0, x_1, ..., x_n = 1\}$ be a partition of $[-1, 1]$ such that $0 \in \Delta x_r$

Let $t_i \in \Delta x_i$, when $i = 1, 2, 3, ..., n$. Now

$$S(P, \beta_2, \beta_1) = \sum_{i=1}^{n} \beta_2(t_i)\{\beta_1(x_i) - \beta_1(x_{i-1})\}$$

$$= \beta_2(t_r)$$

$\therefore \lim_{\mu(P)\to 0} S(P, \beta_2, \beta_1) = 0$ or 1, according as $t_r < 0$ or ≥ 0

Thus, $\lim S(P, \beta_2, \beta_1)$ does not exist.

Let $P^* = P \cup \{0\}$, and $0 \in \Delta x_r$.

Now

$$U(P^*, \beta_2, \beta_1) = 1 \cdot \{\beta_1(x_r) - \beta_1(0)\} = 1$$
$$L(P^*, \beta_2, \beta_1) = 1 \cdot \{\beta_1(x_r) - \beta_1(0)\} = 1$$

Thus,

$$U(P^*, \beta_2, \beta_1) = L(P^*, \beta_2, \beta_1) = 1$$

$\therefore$ $$\beta_2 \in R(\beta_1) \text{ and } \int_{-1}^{1} \beta_2 \, d\beta_1 = 1$$

Ex. 1. For the functions f, β_1, β_2 defined in Example 2, prove that

(*a*) $f \in R(\beta_1)$ iff $f(0+) = f(0)$

and in that case

$$\int_{-1}^{1} f \, d\beta_1 = f(0)$$

(*b*) $f \in R(\beta_2)$ iff $f(0-) = f(0)$,

and in that case

$$\int_{-1}^{1} f \, d\beta_2 = f(0)$$

Ex. 2. Show that

$$\int_0^4 x \, d([x] - x) = \frac{3}{2}$$

where $[x]$ is the greatest integer not exceeding x.

Ex. 3. Show that

(*i*) $\int_0^x d[t] = [x] \;\; \forall \; x \in R$

(*ii*) $\int_0^4 x \, d[x] = 10$

(*iii*) $\int_0^4 x \, d([x] - x) = 2$

(*iv*) $\int_0^2 x^2 \, d(x^2) = 8$

(*v*) $\int_0^2 [x] d(x^2) = 3$

(*vi*) $\int_0^3 x^2 \, d([x] - x) = 5$

(*vii*) $\int_{-1}^{1} (x^2 + e^x) \, d(\text{sgn } x) = 1$

(*viii*) $\int_{\pi}^{2\pi} \sin x \, d(\cos x) = \dfrac{-\pi}{2}$

Ex. 4. Let $\alpha(x) = f(x) = \begin{cases} 0 & \text{if } 0 \le x < 1 \\ 1 & \text{if } 1 \le x \le 2 \end{cases}$

and $$g(x) = \begin{cases} 0 & \text{if } 0 \le x \le 1 \\ 1 & \text{if } 1 < x \le 2 \end{cases}$$

(*i*) Is $f \in R(\alpha)$? If so, compute $\int_0^2 f\, d(\alpha)$.

(*ii*) Is $g \in R(\alpha)$? If so, compute $\int_0^2 g\, d(\alpha)$.

Ex. 5. Evaluate

(*i*) $\int_0^2 x\, d\alpha(x)$, where $\alpha(x) = \begin{cases} x, & 0 \le x \le 1 \\ 2 + x, & 1 < x \le 2 \end{cases}$

(*ii*) $\int_0^3 f(x)\, d([x] + x)$, where $f(x) = \begin{cases} [x], & 0 \le x < 3/2 \\ e^x, & 3/2 \le x \le 3 \end{cases}$

5. SOME IMPORTANT THEOREMS

We add a few theorems before closing the discussion.

Theorem 7. *If $f \in R[a, b]$ and α is monotone increasing on $[a, b]$ such that $\alpha' \in R[a, b]$, then $f \in R(\alpha)$, and*

$$\int_a^b f\, d\alpha = \int_a^b f\alpha'\, dx$$

Let $\varepsilon > 0$ be any given number.

Since f is bounded, there exists $M > 0$, such that

$$|f(x)| \le M, \quad \forall\, x \in [a, b]$$

Again since f, $\alpha' \in R[a, b]$, therefore $f\alpha' \in R[a, b]$ and consequently $\exists\, \delta_1 > 0, \delta_2 > 0$ such that

$$\left| \Sigma f(t_i)\alpha'(t_i)\Delta x_i - \int f\alpha'\, dx \right| < e/2 \qquad \text{...(1)}$$

for $\mu(P) < \delta_1$ and all $t_i \in \Delta x_i$, and

$$\left| \Sigma \alpha'(t_i)\Delta x_i - \int \alpha'\, dx \right| < \varepsilon/4M \qquad \text{...(2)}$$

for $\mu(P) < \delta_2$ and all $t_i \in \Delta x_i$.

Now for $\mu(P) < \delta_2$ and all $t_i \in \Delta x_i$, $s_i \in \Delta x_i$, (2) gives

$$\Sigma|\alpha'(t_i) - \alpha'(s_i)|\,\Delta x_i < 2 \cdot \frac{\varepsilon}{4M} = \frac{\varepsilon}{2M} \qquad ...(3)$$

Let $\delta = \min(\delta_1, \delta_2)$, and P be any partition with $\mu(P) < \delta$.

Then, for all $t_i \in \Delta x_i$, by Lagrange's Mean Value Theorem, there are points $s_i \in \Delta x_i$ such that

$$\Delta\alpha_i = \alpha'(s_i)\Delta x_i \qquad ...(4)$$

Thus

$$\begin{aligned}\left|\Sigma f(t_i)\,\Delta\alpha_i - \int f\alpha'\,dx\right| &= \left|\Sigma f(t_i)\alpha'(s_i)\Delta x_i - \int f\alpha'\,dx\right| \\ &= \left|\Sigma f(t_i)\alpha'(t_i)\Delta x_i - \int f\alpha'\,dx + \Sigma f(t_i)[\alpha'(s_i) - \alpha'(t_i)]\Delta x_i\right| \\ &\le \left|\Sigma f(t_i)\alpha'(t_i)\Delta x_i - \int f\alpha'\,dx\right| + \Sigma|f(t_i)|\,|\alpha'(s_i) - a'(t_i)|\,\Delta x_i \\ &< \frac{\varepsilon}{2} + M\frac{\varepsilon}{2M} = \varepsilon \qquad ...(5)\end{aligned}$$

Hence for any $\varepsilon > 0, \exists\, \delta > 0$ such that for all partitions with $\mu(P) < \delta$, (5) holds

$$\Rightarrow \qquad \lim_{\mu(P)\to 0} \Sigma f(t_i)\,\Delta\alpha_i \text{ exists and equals } \int_a^b f\alpha'\,dx$$

$$\Rightarrow \qquad f \in R(\alpha), \text{ and } \int_a^b f\,d\alpha = \int_a^b f\alpha'\,dx$$

Theorem 8. ***A particular case.*** *If f is continuous on $[a, b]$ and α has a continuous derivative on $[a, b]$, then*

$$\int_a^b f\,d\alpha = \int_a^b f\alpha'\,dx$$

Under the given conditions all the integrals exist.

Let $P = \{a = x_0, ..., x_n = b\}$ be any partition of $[a, b]$. Thus, by Lagrange's Mean Value Theorem it is possible to find $t_i \in\,]x_{i-1}, x_i[$, such that

$$\alpha(x_i) - \alpha(x_{i-1}) = \alpha'(t_i)\,(x_i - x_{i-1});\ i = 1, 2, ..., n$$

or

$$\Delta\alpha_i = \alpha'(t_i)\,\Delta x_i$$

$$\begin{aligned}\therefore \qquad S(P, f, \alpha) &= \sum_{i=1}^{n} f(t_i)\,\Delta\alpha_i \\ &= \sum_{i=1}^{n} f(t_i)\,\alpha'(t_i)\,\Delta x_i = S(P, f\alpha') \qquad ...(6)\end{aligned}$$

Proceeding to limits as $\mu(P) \to 0$, since both the limits exist, we get

$$\int_a^b f\,d\alpha = \int_a^b f\alpha'\,dx$$

Notes:

1. The theorem illustrates one of the situations in which Reimann-Stieltjes integrals reduce to Riemann integrals.
2. In equation (6) lim $S(P, f, \alpha)$ exists in view of Theorem 5 while lim $S(P, f\alpha')$ exists because $f\alpha'$ is continuous and hence integrable in the Riemann sense.

Examples:

(*i*) $$\int_0^2 x^2\,dx^2 = \int_0^2 x^2\,2x\,dx = \int_0^2 2x^3\,dx = 8$$

(*ii*) $$\int_0^2 [x]dx^2 = \int_0^2 [x]2x\,dx$$

$$= \int_0^1 [x]2x\,dx + \int_1^2 [x]2x\,dx = 0 + 3 = 3$$

Ex. Evaluate the following integrals:

(*i*) $\int_1^4 (x - [x])\,dx^2$ (*ii*) $\int_0^3 \sqrt{x}\,dx^3$

(*iii*) $\int_0^3 [x]\,d(e^x)$ (*iv*) $\int_0^{\pi/2} x\,d(\sin x)$

Theorem 9. ***First Mean Value Theorem.*** *If a function f is continuous on* $[a, b]$ *and* α *is monotonic increasing on* $[a, b]$*, then there exists a number* ξ *in* $[a, b]$ *such that*

$$\int_a^b f\,d\alpha = f(\xi)\,\{\alpha(b) - \alpha(a)\}$$

f is continuous and α is monotonic, therefore $f \in R(\alpha)$.

Let m, M be the infimum and supremum of f in $[a, b]$. Then as in § 10.1.1

$$m\{\alpha(b) - \alpha(a)\} \le \int_a^b f\,d\alpha \le M\{\alpha(b) - \alpha(a)\}$$

Hence, there exists a number μ, $m \le \mu \le M$ such that

$$\int_a^b f\,d\alpha = \mu\{\alpha(b) - \alpha(a)\}$$

Again, since f is continuous, there exists a number $\xi \in [a, b]$ such that $f(\xi) = \mu$

$$\therefore \qquad \int_a^b f\,d\alpha = f(\xi)\,\{\alpha(b) - \alpha(a)\}$$

Remark: It may not be possible always to choose ξ such that $a < \xi < b$.

Consider $\alpha(x) = \begin{cases} 0, & x = a \\ 1, & a < x \le b \end{cases}$

For a continuous function f, we have

$$\int_a^b f\,d\alpha = f(a) = f(a)\,\{\alpha(b) - \alpha(a)\}$$

Theorem 10. *If f is continuous and α is monotone on $[a, b]$, then*

$$\int_a^b f\,d\alpha = [f(x)\alpha(x)]_a^b - \int_a^b \alpha\,df$$

Under the given conditions all the integrals exist by Theorem 10.5.

Let $P = \{a = x_0, x_1, ..., x_n = b\}$ be a partition of $[a, b]$.

Choose $t_1, t_2, ..., t_n$ such that $x_{i-1} \le t_i \le x_i$, and let $t_0 = a$, $t_{n+1} = b$, so that $t_{i-1} \le x_{i-1} \le t_i$.

Clearly $Q = \{a, = t_0, t_1, t_2, ..., t_n, t_{n+1} = b\}$ is also a partition of $[a, b]$.

Now

$$\begin{aligned} S(P, f, \alpha) &= \sum_{i=1}^{n} f(t_i)\Delta\alpha_i \\ &= f(t_1)[\alpha(x_1) - \alpha(x_0)] + f(t_2)[\alpha(x_2) - \alpha(x_1)] + ... + f(t_n)[\alpha(x_n) - \alpha(x_{n-1})] \\ &= -\alpha(x_0)\,f(t_1) - \alpha(x_1)[f(t_2) - f(t_1)] + a(x_2)[f(t_3) - f(t_2)] + ... \\ &\qquad + \alpha(x_{n-1})[f(t_n) - f(t_{n-1})] + \alpha(x_n)\,f(t_n) \end{aligned}$$

Adding and subtracting $\alpha(x_0)\,f(t_0) + \alpha(x_n)\,f(t_{n+1})$, we get

$$\begin{aligned} S(P, f, \alpha) &= \alpha(x_n)\,f(t_{n+1}) - \alpha(x_0)\,f(t_0) - \sum_{i=0}^{n} \alpha(x_i)\{f(t_{i+1}) - f(t_i)\} \\ &= f(b)\alpha(b) - f(a)\,\alpha(a) - S(Q, \alpha, f) \qquad ...(1) \end{aligned}$$

If $\mu(P) \to 0$, then $\mu(Q) \to 0$ and Theorem 10.5 shows that lim $S(P, f, \alpha)$ and lim $S(Q, \alpha, f)$ both exist and that

$$\lim S(P, f, \alpha) = \int_a^b f\,d\alpha$$

and $$\lim S(Q, \alpha, f) = \int_a^b \alpha\, df$$

Hence proceeding to limits when $\mu(P) \to 0$, we get from equation (1),

$$\int_a^b f\, d\alpha = [f(x)\, \alpha(x)]_a^b - \int_a^b \alpha\, df \qquad \text{...(2)}$$

where $[f(x)\, \alpha(x)]_a^b$ denotes the difference $f(b)\alpha(b) - f(a)\alpha(a)$.

Remark: The theorem holds when one of the functions is continuous and the other monotone.

Note: The theorem is similar to the theorem, '*Integration by parts*' for Riemann integration.

Corollary. The result of the theorem can be put in a slightly different form, by using Theorem 9, if, in addition to monotonicity α is continuous also

$$\begin{aligned}\int_a^b f\, d\alpha &= f(b)\alpha(b) - f(a)\alpha(a) - \int_a^b \alpha\, df \\ &= f(b)\alpha(b) - f(a)\alpha(a) - a(\xi)\,[f(b) - f(a)] \\ &= f(b)[\alpha(\xi) - \alpha(a)] + f(b)\,[\alpha(b) - \alpha(\xi)]\end{aligned}$$

where $\xi \in [a, b]$.

Stated in this form, it is called the **Second Mean Value Theorem.**

Theorem 11. ***Change of variable.*** *If*

(*i*) *f is a continuous function on $[a, b]$, and*

(*ii*) *ϕ is a continuous and strictly monotonic function on $[\alpha, \beta]$ where $a = \phi(\alpha)$, $b = \phi(\beta)$*

then

$$\int_a^b f(x)\, dx = \int_\alpha^\beta f(\phi(y))\, d\phi(y)$$

$$\left[\text{Change of variable in } \int_a^b f(x)\, dx \text{ by putting } x = \phi(y)\right]$$

Let ϕ be strictly monotonic increasing function.

Since ϕ is strictly monotonic, it is invertible, *i.e.*,

$$x = \phi(y) \;\Rightarrow\; y = \phi^{-1}(x),\; \forall\, x \in [a, b]$$

so that

$$\alpha = \phi^{-1}(a),\; \beta = \phi^{-1}(b)$$

Let

$$P = \{a = x_0, x_1, x_2, ..., x_n = b\}$$

be any partition of $[a, b]$, and

$$Q = (\alpha = y_0, y_1, y_2, ..., y_n = \beta),\ y_i = \phi^{-1}(x_i)$$

be the corresponding partition of $[\alpha, \beta]$, so that

$$\Delta x_i = x_i - x_{i-1} = \phi(y_i), \phi(y_{i-1}) = \Delta\phi_i \qquad ...(1)$$

Again, for any $\xi_i \in \Delta x_i$, let $\eta_i \in \Delta y_i$ where

$$\xi_i = \phi(\eta_i) \qquad ...(2)$$

Putting $g(y) = f(\phi(y))$, we have

$$\begin{aligned} S(P, f) &= \sum_i f(\xi_i)\,\Delta x_i \\ &= \sum_i f(\phi(\eta_i))\,\Delta\phi_i = \sum_i g(\eta_i)\,\Delta\phi_i \\ &= S(Q, g, \phi) \qquad ...(3) \end{aligned}$$

Continuity of f implies that $S(P, f) \to \int_a^b f\,dx$ as $\mu(P) \to 0$. Also continuity of g implies (by Theorem 10.5) that $S(Q, g, \phi) \to \int_\alpha^\beta g(y)\,d\phi$ as $\mu(P) \to 0$.

Since uniform continuity of ϕ on $[\alpha, \beta]$ implies that $\mu(Q) \to 0$ as $\mu(P) \to 0$, therefore letting $\mu(P) \to 0$ in (3), we get

$$\int_a^b f(x)\,dx = \int_\alpha^\beta g(y)\,d\phi = \int_\alpha^\beta f(\phi(y))\,d\phi(y).$$

11 Improper Integrals

1. INTRODUCTION

The concept of Riemann integrals as developed in Chapter 9 requires that the range of integration is finite and the integrand remains bounded in that domain. If either (or both) of these assumptions is not satisfied it is necessary to attach a new interpretation to the integral.

In case the integrand f becomes infinite in the interval $a \le x \le b$, *i.e.*, f has points of *infinite discontinuity* (*singular points*) in $[a, b]$, or the limits of integration a or b (or both) become infinite, the symbol $\int_a^b f\,dx$ is called an *improper* (or *infinite* or *generalised*) integral. Thus

$$\int_1^\infty \frac{dx}{x^2},\ \int_{-\infty}^\infty \frac{dx}{1+x^2},\ \int_0^1 \frac{dx}{x(1-x)},\ \int_{-1}^\infty \frac{dx}{x^2}$$

are examples of improper integrals.

For the sake of distinction, the integrals (of Chapter 9) which are not improper are called *proper* integrals. Thus $\int_0^1 \frac{\sin x}{x}\,dx$ is a proper integral.

It will be assumed throughout that the number of singular points in any interval is *finite* and, therefore, when the range of integration is infinite, that all the singular points can be included in a finite interval. The restriction on the number is not necessary for the existence of the improper integral, but consideration of the discussion is beyond our limits.

Further, it is assumed once for all that *in a finite interval which encloses no point of infinite discontinuity* (*singular point*) *the integrand is bounded and integrable.*

2. INTEGRATION OF UNBOUNDED FUNCTIONS WITH FINITE LIMITS OF INTEGRATION

Definitions. Let a function f be defined in an interval $[a, b]$ everywhere except possibly at a finite number of points.

(*i*) Convergence at the left-end. Let a be the only point of infinite discontinuity of f so that according to the assumption made in the last section, the integral $\int_{a+\lambda}^b f\,dx$ exists for every $\lambda, 0 < \lambda < b-a$.

The improper integral $\int_a^b f\,dx$ is defined as the limit of $\int_{a+\lambda}^b f\,dx$ when $\lambda \to 0+$, so that

$$\int_a^b f\,dx = \lim_{\lambda \to 0+} \int_{a+\lambda}^b f\,dx \qquad \ldots(1)$$

If this limit exists and is finite, the improper integral $\int_a^b f\,dx$ is said to *exist* or *converge* (at a), if otherwise, it is called *divergent*.

Note: For any value c between a and b, $a < c < b$,

$$\int_a^b f\,dx = \int_a^c f\,dx + \int_c^b f\,dx$$

$\int_c^b f\,dx$ is a proper integral, so that the two integrals $\int_a^b f\,dx$ and $\int_a^c f\,dx$ converge and diverge together. Thus while testing the integral $\int_a^b f\,dx$ for convergence at a, it may be replaced by $\int_a^c f\,dx$ for any convenient c such that $a < c < b$.

***(ii)* Convergence at the right-end.** Let b be the only point of infinite discontinuity of f, the improper integral is then defined by the relation

$$\int_a^b f\,dx = \lim_{\mu \to 0+} \int_a^{b-\mu} f\,dx,\ 0 < \mu < b - a \qquad \ldots(2)$$

If the limit exists, the improper integral is said to be *convergent* (at b), otherwise it is called *divergent*.

Note: For the same reason as above, when testing the integral $\int_a^b f\,dx$ for convergence at b, it may be replaced by $\int_c^b f\,dx$ for any convenient c between a and b.

***(iii)* Convergence at both the end-points.** If the end-points a and b are the only points of infinite discontinuity of f, then for any point c within the interval $[a, b]$, the improper integral $\int_a^b f\,dx$ is understood to mean

$$\int_a^c f\,dx + \int_c^b f\,dx \qquad \ldots(3)$$

If *both* the integrals in (3) exist in accordance with the definitions given above, the improper integral *converges*; otherwise it is *divergent*.

The improper integral is also defined as

$$\int_a^b f\,dx = \lim_{\substack{\lambda \to 0+ \\ \mu \to 0+}} \int_{a+\lambda}^{b-\mu} f\,dx \qquad \text{...(3A)}$$

The improper integral exists if the limit exists.

(iv) **Convergence at interior points.** If an interior point c, $a < c < b$, is the only point of infinite discontinuity of f, we put

$$\int_a^b f\,dx = \int_a^c f\,dx + \int_c^b f\,dx \qquad \text{...(4)}$$

The improper integral $\int_a^b f\,dx$ is convergent if both the integrals on the R.H.S. exist in accordance with the definitions given above.

Similarly if the function has a finite number of points of infinite discontinuity, $c_1, c_2, ..., c_m$ within $[a, b]$, where

$$a \le c_1 < c_2 < ... < c_m \le b$$

the improper integral $\int_a^b f\,dx$ is defined as

$$\int_a^b f\,dx = \int_a^{c_1} f\,dx + \int_{c_1}^{c_2} f\,dx + ... + \int_{c_m}^b f\,dx \qquad \text{...(5)}$$

and is said to be convergent if *all* the integrals on the R.H.S. of equation (5) are convergent, otherwise it is divergent.

Example 1. Examine the convergence of

(i) $\displaystyle\int_0^1 \frac{dx}{x^2}$ *(ii)* $\displaystyle\int_0^1 \frac{dx}{\sqrt{1-x}}$ *(iii)* $\displaystyle\int_0^2 \frac{dx}{\left(2x - x^2\right)}$

■ *(i)* 0 is the only point of infinite discontinuity of the integrand in [0, 1]. Thus,

$$\int_0^1 \frac{dx}{x^2} = \lim_{\lambda \to 0+} \int_\lambda^1 \frac{dx}{x^2},\ 0 < \lambda < 1$$

$$= \lim_{\lambda \to 0+} \left(\frac{1}{\lambda} - 1\right) = \infty$$

Thus the improper integral is divergent.

(*ii*) Since 1 is the only point of infinite discontinuity of the integrand in [0, 1], we put

$$\int_0^1 \frac{dx}{\sqrt{1-x}} = \lim_{\mu\to 0+} \int_0^{1-\mu} \frac{dx}{\sqrt{1-x}},\ 0<\mu<1$$

$$= \lim_{\mu\to 0+} \left[-2\sqrt{1-x}\right]_0^{1-\mu}$$

$$= -\lim_{\mu\to 0+} 2\left(\sqrt{\mu}-1\right) = 2$$

Thus the improper integral exists and is equal to 2.

(*iii*) Both the end-points 0, 2 are the points of infinite discontinuity of the integrand and are in fact the only such points in [0, 2].

Thus for any point, say 1, within [0, 2], we put

$$\int_0^2 \frac{dx}{2x-x^2} = \lim_{\lambda\to 0+} \int_\lambda^1 \frac{dx}{x(2-x)} + \lim_{\mu\to 0+} \int_1^{2-\mu} \frac{dx}{x(2-x)}$$

$$= \frac{1}{2}\lim_{\lambda\to 0+}\left[\log\frac{x}{2-x}\right]_\lambda^1 + \frac{1}{2}\lim_{\mu\to 0+}\left[\log\frac{x}{2-x}\right]_1^{2-\mu}$$

$$= -\frac{1}{2}\lim_{\lambda\to 0+}\log\frac{\lambda}{2-\lambda} + \frac{1}{2}\lim_{\mu\to 0+}\log\frac{2-\mu}{\mu}$$

$$= \infty$$

Thus, the given integral diverges.

EXERCISE

1. Test for convergence the improper integrals:

(*i*) $\displaystyle\int_0^1 \frac{dx}{\sqrt{1-x^2}}$ (*ii*) $\displaystyle\int_1^2 \frac{x\,dx}{\sqrt{x-1}}$

(*iii*) $\displaystyle\int_{-2}^2 \frac{dx}{(1-x)x^2}$ (*iv*) $\displaystyle\int_0^\pi \frac{dx}{\sin x}$

(*v*) $\displaystyle\int_0^1 \log x\, dx$ (*vi*) $\displaystyle\int_a^b \frac{dx}{(b-x)^n}$

2. Compute the integrals or prove their divergence:

(*i*) $\displaystyle\int_0^2 \frac{dx}{x^2-4x+3}$ (*ii*) $\displaystyle\int_0^{1/e} \frac{dx}{x(\log x)^2}$

(*iii*) $\int_1^2 \frac{dx}{x\log x}$ (*iv*) $\int_3^5 \frac{x^2\,dx}{\sqrt{(x-3)(5-x)}}$

(*v*) $\int_{-1}^1 \frac{dx}{(2-x)\sqrt{1-x^2}}$ (*vi*) $\int_{-1}^1 \frac{x-1}{x^{5/3}}\,dx$

ANSWERS

1. (*i*) C to $\pi/2$ (*ii*) C to 8/3 (*iii*) D (*iv*) D (*v*) C to -1 (*vi*) C for only $n < 1$.
2. (*i*) D (*ii*) C to 1 (*iii*) D (*iv*) $33\pi/2$ (*v*) $\pi/\sqrt{3}$ (*vi*) D.

3. COMPARISON TESTS FOR CONVERGENCE AT *a* OF $\int_a^b f dx$

(*Integrand retaining its sign*)

Let a, the left end of the interval, be the only point of infinite discontinuity of f in $[a, b]$. The case when b is the only point of infinite discontinuity can be dealt within the same way.

When the integrand keeps the same sign, positive or negative, in a small neighbourhood of a, we may suppose that f is non-negative therein, for, if negative it can be replaced by $(-f)$, for testing the convergence of $\int_a^b f\,dx$. The case $f = 0$ being trivial, there is no loss of generality to suppose that f is *positive* throughout.

The case when f does not necessarily keep the same sign, will be considered in § 3.5 wherein a *general test for convergence* is considered, which holds whether or not f retains its sign.

Theorem 1. *A necessary and sufficient condition for the convergence of the improper integral* $\int_a^b f\,dx$ *at a, where f is positive in* $[a, b]$ *is that there exists a positive number M, independent of* λ, *such that*

$$\int_{a+\lambda}^{b} f\,dx < M, \; 0 < \lambda < b - a$$

We know that the improper integral $\int_a^b f\,dx$ converges at a if for $0 < \lambda < b - a$, $\int_{a+\lambda}^{b} f\,dx$ tends to a finite limit as $\lambda \to 0+$.

Since f is positive in $[a, b]$, the positive function of λ, $\int_{a+\lambda}^{b} f\,dx$ is monotone increasing as λ decreases and will therefore tend to a finite limit if and only if it is bounded above, *i.e.*, there exists a positive number M independent of λ such that

$$\int_{a+\lambda}^{b} f\,dx < M,\ 0 < \lambda < b - a$$

Hence, the proof.

Note: If no such number M exists, the monotonic increasing function $\int_{a+\lambda}^{b} f\,dx$ is not bounded above, and therefore tends to $+\infty$ as $\lambda \to 0+$, and so the improper integral diverges to $+\infty$.

3.1 Comparison Test I (*Comparison of two integrals*)

If f and g be two positive functions such that $f(x) \le g(x)$, for all x in $[a, b]$, then

(*i*) $\int_{a}^{b} f\,dx$ *converges, if* $\int_{a}^{b} g\,dx$ *converges, and*

(*ii*) $\int_{a}^{b} g\,dx$ *diverges, if* $\int_{a}^{b} f\,dx$ *diverges.*

Let f and g be both bounded and integrable in $[a + \lambda, b]$, $0 < \lambda < b - a$ and a is the only point of infinite discontinuity in $[a, b]$.

Since f and g are positive and

$$f(x) \le g(x),\ \ \forall\, x \in [a, b]$$

$$\therefore \qquad \int_{a+\lambda}^{b} f\,dx \le \int_{a+\lambda}^{b} g\,dx \qquad \text{...(1)}$$

(*i*) Let $\int_{a}^{b} g\,dx$ be convergent, so that there exists a positive number M such that

$$\int_{a+\lambda}^{b} g\,dx < M,\ \text{for } 0 < \lambda < b - a$$

Thus, from equation (1)

$$\int_{a+\lambda}^{b} f\,dx < M \text{, for } 0 < \lambda < b - a$$

Hence, $$\int_{a}^{b} f\,dx \text{ converges at } a.$$

(*ii*) Again, if $\int_a^b f\,dx$ is divergent at a, then the positive function $\int_{a+\lambda}^{b} f\,dx$ is not bounded above and therefore from equation (1), $\int_{a+\lambda}^{b} g\,dx$ is also not bounded above. Hence, $\int_a^b g\,dx$ is divergent at a.

3.2 Comparison Test II (*Limit form*)

If f and g are two positive functions in $[a, b]$ such that $\lim\limits_{x \to a+0} \dfrac{f(x)}{g(x)} = l$, where l is a non-zero finite number, then the two integrals $\int_a^b f\,dx$ and $\int_a^b g\,dx$ converge and diverge together at a.

Evidently $l > 0$.

Let ε be a positive number such that $l - \varepsilon > 0$.

Since $\lim\limits_{x \to a+} \dfrac{f(x)}{g(x)} = l$, therefore there exists a neighbourhood $]a, c[$, $a < c < b$, such that for all $x \in]a, c[$

$$\left| \frac{f(x)}{g(x)} - l \right| < \varepsilon$$

or $$(l - \varepsilon) g(x) < f(x) < (l + \varepsilon) g(x) \qquad \text{...(2)}$$

Now,

$$(l - \varepsilon) g(x) < f(x), \;\; \forall\, x \in]a, c[$$

So if $\int_a^b f\,dx$, *i.e.*, if $\int_a^c f\,dx$ converges at a then by comparison test I, $(l - \varepsilon)\int_a^c g\,dx$ and consequently $\int_a^b g\,dx$ converges at a.

Again from equation (2),

$$f(x) < (l+\varepsilon)\, g(x), \quad \forall\, x \in\,]a, c[\,.$$

So if $\int_a^b f\,dx$ *i.e.*, if $\int_a^c f\,dx$ diverges at a, then by comparison test, $(l+\varepsilon)\int_a^c g\,dx$ and therefore, $\int_a^b g\,dx$ diverges at a.

It may similarly be shown that $\int_a^b f\,dx$ converges and diverges with $\int_a^b g\,dx$.

Hence, the two integrals behave alike.

Note: It can be easily shown that:

(*i*) if $f/g \to 0$ and $\int_a^b g\,dx$ converges, then $\int_a^b f\,dx$ also converges, and

(*ii*) if $f/g \to \infty$ and $\int_a^b g\,dx$ diverges, then $\int_a^b f\,dx$ also diverges.

3.3 A Useful Comparison Integral

The improper integral $\int_a^b \frac{dx}{(x-a)^n}$ *converges if and only if* $n<1$.

It is a proper integral if $n \le 0$, and improper for other values of n; a being the only point of infinite discontinuity of the integrand.

Now for $n \ne 1$,

$$\int_a^b \frac{dx}{(x-a)^n} = \lim_{\lambda\to 0+}\int_{a+\lambda}^b \frac{dx}{(x-a)^n}, \quad 0<\lambda<b-a$$

$$= \lim_{\lambda\to 0+}\frac{1}{1-n}\left[\frac{1}{(b-a)^{n-1}} - \frac{1}{\lambda^{n-1}}\right]$$

$$= \begin{cases} 1/(1-n)(b-a)^{n-1}, & \text{if } n<1 \\ \infty, & \text{if } n>1 \end{cases}$$

Again, for $n = 1$

$$\int_a^b \frac{dx}{(x-a)^n} = \int_a^b \frac{dx}{x-a} = \lim_{\lambda \to 0+} \int_{a+\lambda}^b \frac{dx}{x-a}$$

$$= \lim_{\lambda \to 0+} \{\log(b-a) - \log \lambda\} = \infty$$

Thus, $\displaystyle\int_a^b \frac{dx}{(x-a)^n}$ converges only for $n < 1$

The integral is widely used when applying the comparison tests for testing the convergence of improper integrals.

Note: A similar result holds for convergence of $\displaystyle\int_a^b \frac{dx}{(b-x)^n}$ at b.

Example 2. Test the convergence of

(*i*) $\displaystyle\int_0^1 \frac{dx}{\sqrt{1-x^3}}$ (*ii*) $\displaystyle\int_0^{\pi/2} \frac{\sin x}{x^p}\, dx$

■ (*i*) Let $f(x) = \dfrac{1}{\sqrt{1-x^3}} = \dfrac{1}{(1-x)^{1/2}} \dfrac{1}{\left(1+x+x^2\right)^{1/2}}$.

Clearly, $\dfrac{1}{\left(1+x+x^2\right)^{1/2}}$ is a bounded function and let M be its upper bound.

$$\therefore \qquad f(x) \le \frac{M}{(1-x)^{1/2}}$$

Also $\displaystyle\int_0^1 \frac{dx}{(1-x)^{1/2}}$ is convergent.

Therefore by comparison test, $\displaystyle\int_0^1 \frac{dx}{\sqrt{1-x^3}}$ is convergent.

(*ii*) For $p \le 1$, it is a proper integral. For $p > 1$, it is an improper integral, 0 being the point of infinite discontinuity.

Now

$$\frac{\sin x}{x^p} = \frac{1}{x^{p-1}} \cdot \frac{\sin x}{x}$$

The function $\frac{\sin x}{x}$ is bounded and $\frac{\sin x}{x} \leq 1$.

$$\therefore \qquad \frac{\sin x}{x^p} \leq \frac{1}{x^{p-1}}$$

Also $\int_0^{\pi/2} \frac{dx}{x^p - 1}$ converges only if $p - 1 < 1$ or $p < 2$.

Therefore by comparison test, $\int_0^{\pi/2} \frac{\sin x}{x^p}\, dx$ converges for $p < 2$ and diverges for $p \geq 2$.

3.4 With the help of § 3.1, 3.2 and 3.3, we deduce two comparison tests which are of much practical utility.

I. *If f is positive in a nbd of a, then the integral $\int_a^b f\, dx$ converges at a if there exists a positive number n less than 1 and a fixed positive number M such that $f(x) \leq M/(x-a)^n$ for all x in $]a, b]$.*

Also, $\int_a^b f\, dx$ diverges if there exists a number $n \geq 1$ and a fixed positive number G such that

$$f(x) \geq G/(x-a)^n \text{ in }]a, b].$$

II. *If $\lim_{x \to a+} \left[(x-a)^n f(x)\right]$ exists and is non-zero finite, then the integral $\int_a^b f\, dx$ converges if and only if $n < 1$.*

Example 3. Find the values of m and n for which the following integrals converge:

(*i*) $\int_0^1 e^{-mx} x^n \, dx$

(*ii*) $\int_0^1 (\log 1/x)^m \, dx$

■ (*i*) Let k be a number greater than 1 and e^{-m}, for all m.

In [0, 1], $e^{-mx}x^n \le kx^n$, for all m, and $\int_0^1 x^n\,dx = \int_0^1 \frac{dx}{x^{-n}}$ converges for $-n < 1$ or $n > -1$ only. Thus, $\int_0^1 e^{-mx}x^n\,dx$ converges only for $n > -1$, irrespective of the values of m.

(*ii*) Putting $\int_0^1 (\log 1/x)^m\,dx = \int_0^{1/2} (\log 1/x)^m\,dx + \int_{1/2}^1 (\log 1/x)^m\,dx$, 0 and 1 respectively are the points of infinite discontinuity of the integrals on the right.

Let $f(x) = (\log 1/x)^m$.

Convergence at 0.

$\int_0^{1/2} (\log 1/x)^m\,dx$ is a proper integral if $m \le 0$, for the integrand tends to a finite limit as $x \to 0$ (1 if $m = 0$; 0 if $m < 0$). 0 is the only point of infinite discontinuity if $m > 0$.

For $m > 0$, take

$$g(x) = \frac{1}{x^p},\ 0 < p < 1$$

so that

$$\frac{f(x)}{g(x)} = x^p (\log 1/x)^m \to 0 \text{ as } x \to 0.$$

Also $\int_0^{1/2} g\,dx$ converges, therefore $\int_0^{1/2} (\log 1/x)^m\,dx$ converges. Thus, $\int_0^{1/2} (\log 1/x)^m\,dx$ converges for all m.

Convergence at 1.

$\int_{1/2}^1 (\log 1/x)^m\,dx$ is a proper integral if $m \ge 0$, and 1 is the only point of infinite discontinuity if $m < 0$.

For $m < 0$, let

$$g(x) = \frac{1}{(1-x)^{-m}}$$

so that

$$\frac{f(x)}{g(x)} = \left(\frac{\log 1/x}{1-x}\right)^m \to 1 \text{ as } x \to 1.$$

Hence, the two integrals $\int_{1/2}^{1} f\,dx$ and $\int_{1/2}^{1} g\,dx$ behave alike.

$\int_{1/2}^{1} \frac{dx}{(1-x)^{-m}}$ converges if $-m < 1$ or $m > -1$ therefore $\int_{1/2}^{1} (\log 1/x)^m\,dx$ also converges if $0 > m > -1$.

Consequently $\int_{0}^{1} (\log 1/x)^m\,dx$ is convergent when $0 > m > -1$.

Example 4. Show that

(*i*) $\int_{0}^{1} \frac{\log x}{\sqrt{x}}\,dx$ is convergent, but

(*ii*) $\int_{1}^{2} \frac{\sqrt{x}}{\log x}\,dx$ is divergent.

■ (*i*) Since $\log x/\sqrt{x}$ is negative in [0, 1], we take

$$f(x) = -\frac{\log x}{\sqrt{x}} = \frac{\log 1/x}{\sqrt{x}}$$

0 is the only point of infinite discontinuity.

Let $g(x) = \frac{1}{x^{3/4}}$ so that

$$\frac{f(x)}{g(x)} = x^{1/4}(\log 1/x) \to 0 \text{ as } x \to 0+$$

The integral $\int_{0}^{1} g\,dx = \int_{0}^{1} \frac{dx}{x^{3/4}}$ is convergent at 0; therefore $\int_{0}^{1} \frac{\log 1/x}{\sqrt{x}}\,dx$ and so $\int_{0}^{1} \frac{\log x}{\sqrt{x}}\,dx$ is also convergent.

(*ii*) Here 1 is the only point of infinite discontinuity.

Let $f(x) = \frac{\sqrt{x}}{\log x}$ and $g(x) = \frac{1}{x-1}$, so that

$$\lim_{x\to 1+} \frac{f(x)}{g(x)} = \lim_{x\to 1+} \frac{(x-1)\sqrt{x}}{\log x} = \lim_{x\to 1+} \frac{\frac{3}{2}x^{1/2} - \frac{1}{2}x^{-1/2}}{1/x} = 1$$

The integral $\int_1^2 \frac{dx}{x-1}$ diverges and therefore $\int_1^2 \frac{\sqrt{x}}{\log x}\,dx$ also diverges.

Example 5. Show that the integral $\int_0^{\pi/2} \left(\frac{\sin^m x}{x^n}\right) dx$ exists if and only if $n < m + 1$.

■ Let $f(x) = \frac{\sin^m x}{x^n} = \frac{1}{x^{n-m}} \cdot \left(\frac{\sin x}{x}\right)^m$

Here as $x \to 0+$, $f(x) \to 0$ if $n - m < 0$, and $f(x) \to \infty$ if $n - m > 0$.

Thus, it is a proper integral if $n \leq m$, and improper if $n > m$, 0 being the only point of infinite discontinuity of f.

When $n > m$, let

$$g(x) = \frac{1}{x^{n-m}}$$

so that

$$\frac{f(x)}{g(x)} = \left(\frac{\sin x}{x}\right)^m \to 1 \text{ as } x \to 0+$$

Also $\int_0^{\pi/2} g\,dx = \int_0^{\pi/2} \frac{dx}{x^{n-m}}$ converges if and only if $n - m < 1$ or $n < m + 1$, therefore $\int_0^{\pi/2} \frac{\sin^m x}{x^n} dx$ also converges if and only if $n < m + 1$ which includes the case $n \leq m$ when the integral is proper.

Example 6. Show that $\int_0^1 x^{m-1}(1-x)^{n-1}\,dx$ exists if and only if m, n are both positive.

■ It is a proper integral for $m \geq 1, n \geq 1$, 0 and 1 are the only points of infinite discontinuity; 0 when $m < 1$, and 1 when $n < 1$.

For $m < 1$ and $n < 1$.

Taking a number, say $\frac{1}{2}$, between 0 and 1, we put

$$\int_0^1 x^{m-1}(1-x)^{n-1}\,dx = \int_0^{1/2} x^{m-1}(1-x)^{n-1}\,dx + \int_{1/2}^1 x^{m-1}(1-x)^{n-1}\,dx$$

and examine the two integrals on the right for convergence at 0 and 1 respectively.

Convergence at 0, when $m < 1$.

Let

$$f(x) = x^{m-1}(1-x)^{n-1} = \frac{(1-x)^{n-1}}{x^{1-m}}$$

and
$$g(x)=\frac{1}{x^{1-m}}$$

then
$$\frac{f(x)}{g(x)}=(1-x)^{n-1}\to 1 \text{ as } x\to 0$$

Also $\int_0^{1/2} g\,dx \int_0^{1/2}\frac{dx}{x^{1-m}}$ converges at 0 if and only if $1-m<1$, *i.e.*, $m>0$

Thus $\int_0^{1/2} x^{m-1}(1-x)^{n-1}\,dx$ converges at zero if and only if $m>0$

Convergence at 1, when $n<1$
$$f(x)=x^{m-1}(1-x)^{n-1}=\frac{x^{m-1}}{(1-x)^{1-n}}$$

Let
$$g(x)=\frac{1}{(1-x)^{1-n}}$$

then
$$\frac{f(x)}{g(x)}=x^{m-1}\to 1 \text{ as } x\to 1$$

Also $\int_{1/2}^{1} g\,dx=\int_{1/2}^{1}\frac{dx}{(1-x)^{1-n}}$ converges at 1, if and only if $1-n<1$, or $n>0$.

Thus $\int_{1/2}^{1} x^{m-1}(1-x)^{n-1}\,dx$ converges at 1, if and only if $n>0$.

Hence $\int_0^1 x^{m-1}(1-x)^{n-1}\,dx$ exists for positive values of m and n only.

This integral $\int_0^1 x^{m-1}(1-x)^{n-1}\,dx$, for $m, n>0$ is called *Beta function* and is denoted by $\beta(m, n)$.

Example 7. For what values of m and n is the integral
$$\int_0^1 x^{m-1}(1-x)^{n-1}\log x\,dx$$
convergent?

■ The integrand is negative in [0, 1], therefore we shall test for convergence the integral

$$\int_0^1 -x^{m-1}(1-x)^{n-1}\log x\,dx,\ i.e.,\ \int_0^1 x^{m-1}(1-x)^{n-1}\log\frac{1}{x}\,dx$$

Since 0 and 1 are the possible points of infinite discontinuity, therefore taking a number, say $\frac{1}{2}$, between 0 and 1, we examine the integrals

$$\int_0^{1/2} x^{m-1}(1-x)^{n-1}\log\frac{1}{x}\,dx,\ \int_{1/2}^1 x^{m-1}(1-x)^{n-1}\log\frac{1}{x}\,dx$$

at 0 and 1 respectively.

Convergence at 0.

It is a proper integral for $m-1>0$, and improper for $m \le 1$, 0 being the only point of infinite discontinuity then.

For $m \le 1$, let

$$f(x) = x^{m-1}(1-x)^{n-1}\log\frac{1}{x} = \frac{(1-x)^{n-1}\log 1/x}{x^{1-m}}$$

and

$$g(x) = \frac{1}{x^p}$$

$\int_0^{1/2} g\,dx$ is convergent if and only if $p<1$.

Also $\dfrac{f(x)}{g(x)} = x^{p+m-1}(1-x)^{n-1}\log\dfrac{1}{x} \to 0$ as $x \to 0+$ if $p+m-1>0$ or $m>1-p>0$.

So by comparison test, $\int_0^{1/2} f\,dx$ converges if and only if $m>0$.

Convergence at 1.

Since $\lim\limits_{x\to 1-}\dfrac{\log 1/x}{(1-x)^{1-n}} = \lim\limits_{x\to 1}\dfrac{(1-x)^n}{(1-n)^x}$ exists finitely when $n \ge 0$, therefore the integral is proper for $n \ge 0$ and improper for $n<0$, 1 being the only singular point.

For $n<0$, let

$$f(x) = x^{m-1}(1-x)^{n-1}\log\frac{1}{x} = \frac{x^{m-1}\log 1/x}{(1-x)^{1-n}}$$

and

$$g(x) = \frac{1}{(1-x)^q}$$

$\int_{1/2}^{1} g\,dx$ is convergent if and only if $q < 1$.

Also $\dfrac{f(x)}{g(x)} = \dfrac{x^{m-1}\log 1/x}{(1-x)^{1-n-q}}$, tends to a finite limit as $x \to 1$, if $1-n-q \le 1$, *i.e.*, if $n \ge -q > -1$.

Thus, $\int_{1/2}^{1} f\,dx$ converges at 1 if $n > -1$.

Hence, the given integral is convergent when $m > 0, n > -1$.

Example 8. Show that the integral $\int_{0}^{\pi/2} \log\,\sin x\,dx$ is convergent and hence evaluate it.

■ Let $f(x) = \log \sin x$.

As f is non-positive in $[0, \pi/2]$, we consider the function $(-f)$ for testing convergence of the integral.

0 is the only point of infinite discontinuity of f.

Let $g(x) = \dfrac{1}{x^m}$, $m < 1$, so that

$$\frac{-f(x)}{g(x)} = -x^m \log\,\sin\,x \to 0 \text{ as } x \to 0$$

Also $\int_{0}^{\pi/2} \dfrac{dx}{x^m}$ is convergent for $m < 1$.

Therefore, $\int_{0}^{\pi/2} -\log\,\sin x\,dx$ and so $\int_{0}^{\pi/2} \log\,\sin x\,dx$ is convergent.

To evaluate the integral, let $I = \int_{0}^{\pi/2} \log\,\sin x\,dx$.

We know that

$$\sin 2x = 2\sin x \cos x$$

$$\therefore \qquad \log \sin 2x = \log 2 + \log \sin x + \log \cos x$$

$$\Rightarrow \qquad \int_{0}^{\pi/2} \log \sin 2x\,dx = \int_{0}^{\pi/2} (\log 2)dx + \int_{0}^{\pi/2} \log \sin x\,dx + \int_{0}^{\pi/2} \log \cos x\,dx$$

$$= \frac{\pi}{2}\log 2 + I + \int_0^{\pi/2} \log \cos x \, dx$$

Putting $2x = t$ in the first and $x = \frac{\pi}{2} - y$ in the last integral, we get

$$I = \frac{\pi}{2}\log 2 + I + I$$

$\therefore$ $$I = -\frac{\pi}{2}\log 2.$$

EXERCISE

1. Test for convergence:

(*i*) $\int_0^1 \frac{x^n}{1-x}\,dx$ (*ii*) $\int_0^1 \frac{x^n}{1+x}\,dx$

(*iii*) $\int_0^1 \frac{\sin x}{x^{3/2}}\,dx$ (*iv*) $\int_2^3 \frac{x^2+1}{x^2-4}\,dx$

(*v*) $\int_0^1 x^{m-1} e^{-x}\,dx$ (*vi*) $\int_0^{\pi/2} \sin^{p-1} x \cos^{q-1} x\,dx$

(*vii*) $\int_0^{\pi} \frac{\sqrt{x}}{\sin x}\,dx$ (*viii*) $\int_0^1 \frac{x^n \log x}{(1+x)^2}\,dx$

2. Show that the integral $\int_0^{\pi/2} \sin x \, \log \sin x \, dx$ converges to the value $\log 2 - 1$.

[*Hint*: Integrate by parts the integral $\int_\lambda^{\pi/2} \sin x \, \log \sin x \, dx$ and take the limit.]

3. Show that $\int_0^{\pi/2} \cos 2nx \log \sin x \, dx$, $n \geq 1$ converges to the value $-\pi/4n$.

4. Compute, if possible, the integrals

(*i*) $\int_0^{\pi} x \log \sin x \, dx$ (*ii*) $\int_0^{\pi/2} x \cos x \, dx$

(*iii*) $\int_0^1 \frac{\sin^{-1} x}{x}\,dx$ (*iv*) $\int_0^1 \frac{\log x}{\sqrt{1-x^2}}\,dx$

[*Hint*: (*i*) Put $x = \pi - t$ (*ii*) Integrate by parts

(*iii*) Put $x = \sin t$ (*iv*) Put $x = \sin t$.]

ANSWERS

1. (*i*) Div. (*ii*) Conv. for $n > -1$ (*iii*) Conv.
(*iv*) Div. (*v*) Conv. for $m > 0$ (*vi*) Conv. for $p > 0$ and $q > 0$
(*vii*) Div. (*viii*) Conv. for $n > -1$.

4. (*i*) $-\dfrac{\pi^2}{2}\log 2$ (*ii*) $\dfrac{\pi}{2}\log 2$

(*iii*) $\dfrac{\pi}{2}\log 2$ (*iv*) $-\dfrac{\pi}{2}\log 2$

3.5 General Test for Convergence (*Integrand may change sign*)

We now discuss a general test for convergence of an improper integral (finite limits of integration but discontinuous integrand) which holds whether or not the integrand keeps the same sign.

Cauchy's Test. *The improper integral* $\int_a^b f\,dx$ *converges at a if and only if to every* $\varepsilon > 0$ *there corresponds* $\delta > 0$ *such that*

$$\left|\int_{a+\lambda_1}^{a+\lambda_2} f\,dx\right| < \varepsilon,\ 0 < \lambda_1,\ \lambda_2 < \delta$$

[Notice that $\int_{a+\lambda_1}^{a+\lambda_2} f\,dx$ tends to 0 as $\lambda_1, \lambda_2 \to 0$.]

The improper integral $\int_a^b f\,dx$ is said to exist when $\lim\limits_{\lambda\to 0+}\int_{a+\lambda}^{b} f\,dx$ exists finitely.

Let $F(\lambda) = \int_{a+\lambda}^{b} f\,dx$

so that $F(\lambda)$ is a function of λ.

According to Cauchy's criterion for finite limits (§ 1.3, Ch. 5) $F(\lambda)$ tends to a finite limit as $\lambda \to 0$ if and only if for every $\varepsilon > 0$ there corresponds $\delta > 0$ such that for all positive $\lambda_1, \lambda_2 < \delta$.

$$|F(\lambda_1) - F(\lambda_2)| < \varepsilon$$

i.e.,

$$\left|\int_{a+\lambda_1}^{b} f\,dx - \int_{a+\lambda_2}^{b} f\,dx\right| < \varepsilon$$

or

$$\left| \int_{a+\lambda_1}^{a+\lambda_2} f \, dx \right| < \varepsilon$$

3.6 Absolute Convergence

Definition. The improper integral $\int_a^b f \, dx$ is said to be absolutely convergent if $\int_a^b |f| \, dx$ is convergent.

With the help of Cauchy's test, we now deduce a sufficient condition for the convergence of an improper integral.

Theorem 2. *Every absolutely convergent integral is convergent, or*

$$\int_a^b f \, dx \text{ exists if } \int_a^b |f| \, dx \text{ exists.}$$

Since $\int_a^b |f| dx$ exists, therefore by Cauchy's test, for $\varepsilon > 0$ there corresponds $\delta > 0$ such that

$$\left| \int_{a+\lambda_1}^{a+\lambda_2} |f| dx \right| < \varepsilon, \; 0 < \lambda_1, \lambda_2 < \delta \qquad \text{...(1)}$$

Also, we know (Th. 10, Ch. 9) that

$$\left| \int_{a+\lambda_1}^{a+\lambda_2} f \, dx \right| \le \int_{a+\lambda_1}^{a+\lambda_2} |f| dx \qquad \text{...(2)}$$

Hence, from equations (1) and (2)

$$\left| \int_{a+\lambda_1}^{a+\lambda_2} f \, dx \right| < \varepsilon, \; 0 < \lambda_1, \lambda_2 < \delta$$

$\Rightarrow$ $\int_a^b f \, dx$ exists.

Aliter. Since

$$0 \le |f| - f \le 2|f|,$$

therefore by comparison test

$$\int_a^b \{|f| - f\} \, dx \text{ converges}$$

Hence, $\int_a^b f\,dx = \int_a^b \{f - |f|\}\,dx + \int_a^b |f|\,dx$ converges.

Notes:

1. Since $|f|$ is always positive, comparison tests of § 3 are applicable for examining the convergence of $\int_a^b |f|\,dx$, *i.e.*, absolute convergence of $\int_a^b f\,dx$.
2. Every convergent integral is not absolutely convergent. For this reason, a convergent integral which is not absolutely convergent is called a *conditionally convergent* integral.

Example 9. Show that $\int_0^1 \frac{\sin 1/x}{x^p}\,dx$, $p > 0$, converges absolutely for $p < 1$.

■ Let

$$f(x) = \frac{\sin 1/x}{x^p},\ p > 0$$

0 is the only point of infinite discontinuity, and f does not keep the same sign in any neighbourhood of 0.

In [0, 1]

$$|f(x)| = \left|\frac{\sin 1/x}{x^p}\right| = \frac{|\sin 1/x|}{x^p} < \frac{1}{x^p}$$

Also $\int_0^1 \frac{dx}{x^p}$ converges if and only if $p < 1$.

Hence by comparison test, the integral $\int_0^1 \left|\frac{\sin 1/x}{x^p}\right| dx$ converges and so $\int_0^1 \frac{\sin 1/x}{x^p}\,dx$ converges absolutely if and only if $p < 1$.

Conditional convergence of improper integrals of this type is generally tested by reducing them to the other type, with infinite range of integration. See examples § 11.5.

4. INFINITE RANGE OF INTEGRATION

We shall now consider the convergence of improper integrals of bounded* integrable function with infinite range of integration (a or b or both infinite).

* The word bounded seems to be redundant. But some authors extend the class of functions integrable in the Riemann sense to include those unbounded functions whose improper integrals exists. The word bounded is inserted here to exclude the possibility of interpreting the term integrable functions in the extended sense.

Definitions.

(*i*) Convergence at ∞. The symbol

$$\int_a^\infty f\,dx,\ x \geq a \qquad \ldots(1)$$

is defined as the limit of $\int_a^X f\,dx$, when $X \to \infty$, so that

$$\int_a^\infty f\,dx = \lim_{X\to\infty} \int_a^X f\,dx \qquad \ldots(2)$$

If the limit exists and is finite, then the improper integral (1) is said to be *convergent*, otherwise it is said to be *divergent*.

Note: For $a_1 > a$,

$$\int_a^X f\,dx = \int_a^{a_1} f\,dx + \int_{a_1}^X f\,dx$$

which implies that the integrals $\int_a^\infty f\,dx$ and $\int_{a_1}^\infty f\,dx$ are either both convergent or both divergent. Thus when testing the integral $\int_a^\infty f\,dx$ for convergence, we can replace it by the integral $\int_{a_1}^\infty f\,dx$ for any convenient $a_1 > a$.

Ex. Examine for convergence

$$\int_0^\infty \frac{x\,dx}{1+x^2},\ \int_1^\infty \frac{dx}{\sqrt{x}},\ \int_a^\infty \sin x\,dx$$

(*ii*) Convergence at $-\infty$

$$\int_{-\infty}^b f\,dx,\ x \leq b \qquad \ldots(3)$$

is defined by the equation

$$\int_{-\infty}^b f\,dx = \lim_{X\to-\infty} \int_X^b f\ dx \qquad \ldots(4)$$

If the limit exists and is finite then the integral (3) *converges* (or *exists*). Otherwise it *diverges* (or *does not exist*).

(*iii*) Convergence at both ends

$$\int_{-\infty}^\infty f\,dx\ \ \forall\, x \qquad \ldots(5)$$

is understood to mean

$$\int_{-\infty}^{c} f\,dx + \int_{c}^{\infty} f\,dx \qquad \text{...(6)}$$

where c is any real number.

If *both* the integrals in (6) exist in accordance with the definition given above, then the integral equation (5) converges, otherwise is divergent.

Integral, (5) can also be defined by the relation

$$\int_{-\infty}^{\infty} f\,dx = \lim_{\substack{X \to \infty \\ Y \to -\infty}} \int_{Y}^{X} f\,dx \qquad \text{...(7)}$$

(*iv*) Integrals of unbounded functions with infinite limits of integration.

When the infinite range of integration includes a finite number of points of infinite discontinuity of f, consider an interval $[a, b]$ which contains all the points of discontinuity. The integral $\int_{-\infty}^{\infty} f\,dx$ is then understood to mean

$$\int_{-\infty}^{a} f\,dx + \int_{a}^{b} f\,dx + \int_{b}^{\infty} f\,dx \qquad \text{...(8)}$$

In case *all* the integrals in (8) exist in accordance with the definitions given above, the integral $\int_{-\infty}^{\infty} f\,dx$ converges, otherwise it is divergent.

Example 10. Examine for convergence the integrals:

(*i*) $\displaystyle\int_{0}^{\infty} \sin x\,dx$ (*ii*) $\displaystyle\int_{-\infty}^{\infty} \frac{dx}{1+x^2}$

(*iii*) $\displaystyle\int_{2}^{\infty} \frac{2x^2\,dx}{x^4-1}$ (*iv*) $\displaystyle\int_{-\infty}^{\infty} \frac{dx}{\left(x^2+1\right)^2}$

(*v*) $\displaystyle\int_{0}^{\infty} x^3 e^{-x^2}\,dx$

■ (*i*) By definition

$$\int_{0}^{\infty} \sin x\,dx = \lim_{X \to \infty} \int_{0}^{X} \sin x\,dx = \lim_{X \to \infty} (1 - \cos X)$$

Thus the improper integral does not exist since $\cos X$ has no limit when $X \to \infty$.

(*ii*) $$\int_{-\infty}^{\infty} \frac{dx}{1+x^2} = \lim_{\substack{X\to\infty \\ Y\to-\infty}} \int_{Y}^{X} \frac{dx}{1+x^2}$$

$$= \lim_{\substack{X\to\infty \\ Y\to-\infty}} \left(\tan^{-1} X - \tan^{-1} Y\right) = \frac{\pi}{2} - \left(-\frac{\pi}{2}\right) = \pi$$

Thus, the integral exists and is equal to π.

(*iii*) $$\int_{2}^{\infty} \frac{2x^2}{x^4-1}\,dx = \lim_{X\to\infty} \int_{2}^{X} \frac{2x^2}{x^4-1}\,dx$$

$$= \lim_{X\to\infty} \left[\tan^{-1} X - \tan^{-1} 2 + \frac{1}{2}\log\frac{X-1}{X+1} + \frac{1}{2}\log 3\right]$$

$$= \frac{\pi}{2} - \tan^{-1} 2 + \frac{1}{2}\log 3$$

Thus, the integral converges.

(*iv*) $$\int_{-\infty}^{\infty} \frac{dx}{\left(1+x^2\right)^2} = \int_{-\infty}^{0} \frac{dx}{\left(1+x^2\right)^2} + \int_{0}^{\infty} \frac{dx}{\left(1+x^2\right)^2}$$

$$= 2\int_{0}^{\infty} \frac{dx}{\left(1+x^2\right)^2} = \lim_{X\to\infty} 2\int_{0}^{X} \frac{dx}{\left(1+x^2\right)^2}$$

$$= \lim_{X\to\infty} 2\left[\tan^{-1} X + \frac{X}{1+X^2}\right] \text{ (by putting } x = \tan\theta\text{)}$$

$= \dfrac{\pi}{2}$, so that the integral converges.

(*v*) $$\int_{0}^{\infty} x^3 e^{-x^2}\,dx = \lim_{X\to\infty} \int_{0}^{X} x^3 e^{-x^2}\,dx$$

$$= \lim_{X\to\infty} \left[\frac{1}{2} - \frac{1}{2}\left(X^2+1\right)e^{-X^2}\right]$$

$= \dfrac{1}{2}$, converges.

Ex. Compute the following integrals or prove their divergence:

(*i*) $\displaystyle\int_{1}^{\infty} \frac{dx}{x^2(x+1)}$ (*ii*) $\displaystyle\int_{\sqrt{2}}^{\infty} \frac{dx}{x\sqrt{x^2-1}}$

(iii) $\displaystyle\int_{a^2}^{\infty} \frac{dx}{x\sqrt{1+x^2}}$ (iv) $\displaystyle\int_0^{\infty} x \sin x \, dx$

(v) $\displaystyle\int_0^{\infty} e^{-\sqrt{x}} dx$ (vi) $\displaystyle\int_1^{\infty} \frac{\tan^{-1} x}{x^2} dx$

(vii) $\displaystyle\int_0^{\infty} e^{-ax} \cos bx \, dx$ (viii) $\displaystyle\int_0^{\infty} \frac{dx}{1+x^3}$

(ix) $\displaystyle\int_0^{\infty} x^2 e^{-x} dx$ (x) $\displaystyle\int_1^{\infty} \frac{\sqrt{x}}{(1+x)^2} dx$

ANSWERS

(i) $1 - \log 2$ (ii) $\pi/4$ (iii) $\log \dfrac{\sqrt{a^2+1}+1}{a^2}$ (iv) Div.

(v) 2 (vi) $\dfrac{\pi}{4} + \dfrac{1}{2}\log 2$ (vii) $\dfrac{a}{a^2+b^2}$ if $a > 0$, div. if $a \le 0$

(viii) $\dfrac{2\pi}{3\sqrt{3}}$ (ix) 2 (x) $\dfrac{1}{2} + \dfrac{\pi}{4}$.

4.1 Comparison Tests for Convergence at ∞ (*Integrand retaining its sign*)

By definition

$$\int_a^{\infty} f \, dx = \lim_{X \to \infty} \int_a^X f \, dx$$

where f is assumed to be bounded and integrable in $[a, X]$ for every $X \ge a$.

As in the case of unbounded functions with finite limits of integration, there is no loss of generality in supposing that the integrand f is positive in $[a, X]$, for, if negative it can be replaced by $(-f)$.

Theorem 3. *A necessary and sufficient condition for the convergence of* $\int_a^{\infty} f \, dx$, *where f is positive in $[a, X]$ is that there exists a positive number M, independent of X, such that*

$$\int_a^X f \, dx < M, \text{ for every } X \ge a$$

The integral $\int_a^{\infty} f\,dx$ is said to be convergent if $\int_a^{X} f\,dx$ tends to a finite limit as $X \to \infty$.

Since f is positive in $[a, X]$, the positive function of X, $\int_a^{X} f\,dx$, is monotone increasing as X increases and will therefore tend to a finite limit if and only if it is bounded above, *i.e.*, there exists a positive number M, independent of X such that

$$\int_a^X f\,dx < M, \text{ for every } X \geq a$$

Hence, the proof.

Note: If no such number M exists, the monotonic increasing function $\int_a^X f\,dx$ is non-bounded above and therefore tends to ∞, as $X \to \infty$ and so $\int_a^{\infty} f\,dx$ diverges to ∞.

4.2 Comparison Test I (*Comparison of two integrals*)

If f and g are positive and $f(x) \leq g(x)$, for all x in $[a, X]$, then

(*i*) $\int_a^{\infty} f\,dx$ *converges, if* $\int_a^{\infty} g\,dx$ *converges, and*

(*ii*) $\int_a^{\infty} g\,dx$ *diverges, if* $\int_a^{\infty} f\,dx$ *diverges.*

Let f and g be both bounded and integrable in $[a, X]$, $X \geq a$.

Since f and g are both positive, and

$$f(x) \leq g(x), \quad \forall\, x \in [a, X]$$

$$\therefore \qquad \int_a^X f\,dx \leq \int_a^X g\,dx \qquad \ldots(1)$$

(*i*) Let $\int_a^{\infty} g\,dx$ be convergent so that there exists a positive number M such that

$$\int_a^X g\,dx < M, \quad X \geq a$$

Hence, from equation (1), $\int_a^X f\,dx < M$, $X \geq a$

Hence, $\int_a^\infty f\,dx$ is convergent.

(*ii*) If $\int_a^\infty f\,dx$ is divergent then the positive function $\int_a^X f\,dx$ is not bounded above and therefore in view of (1), $\int_a^X g\,dx$ is also not bounded above.

Hence, $\int_a^\infty g\,dx$ diverges.

Comparison Test II (*Limit form*)

If f and g are positive in $[a, x]$ *and* $\lim_{x\to\infty} \frac{f}{g} = l$, *where l is a non-zero finite number, then the two integrals* $\int_a^\infty f\,dx$ *and* $\int_a^\infty g\,dx$ *converge or diverge together. Also if* $f/g \to 0$ *and* $\int_a^\infty g\,dx$ *converges then* $\int_a^\infty f\,dx$ *converges, and if* $f/g \to \infty$ *and* $\int_a^\infty g\,dx$ *diverges, then* $\int_a^\infty f\,dx$ *diverges.*

Evidently $l > 0$.

Let ε be a positive number such that $l - \varepsilon > 0$.

Since $\lim_{x\to\infty} \frac{f}{g} = l$, therefore there exists a number k (> a), however large, such that for all $x > k$,

$$\left| \frac{f(x)}{g(x)} - l \right| < \varepsilon$$

or

$$(l - \varepsilon)\, g(x) < f(x) < (l + \varepsilon)\, g(x) \qquad \text{...(2)}$$

Now

$$(l - \varepsilon)\, g(x) < f(x), \;\; \forall\, x > k > a$$

so that if $\int_k^\infty f\,dx$ converges then by Comparison test I, $\int_k^\infty g\,dx$ and therefore $\int_a^\infty g\,dx$ converges at ∞.

Again from (2),

$$f(x) < (l + \varepsilon)\, g(x), \;\; \forall\, x > k > a$$

so that if $\int_k^\infty f\,dx$ diverges, then by comparison test, $\int_k^\infty g\,dx$ and therefore $\int_a^\infty g\,dx$ diverges at ∞.

When $f/g \to 0$, we can find k so that

$$\frac{f(x)}{g(x)} < \varepsilon, \ \forall\, x > k$$

i.e.,

$$f(x) < \varepsilon g(x), \ \forall\, x > k$$

so, if $\int_a^\infty g\,dx$ converges, then $\int_a^\infty f\,dx$ also converges.

When $f/g \to \infty$, we can find numbers k, M such that

$$\frac{f(x)}{g(x)} > M \text{ or } f(x) > Mg(x), \ \forall\, x \geq k$$

Hence, if $\int_a^\infty g\,dx$ diverges, then $\int_a^\infty f\,dx$ also diverges.

4.3 A Useful Comparison Integral

Show that the improper integral $\int_a^\infty \frac{C}{x^n}dx$, $a > 0$, *where C is a positive constant, converges if and only if* $n > 1$.

We have

$$\int_a^X \frac{C}{x^n}dx = \begin{cases} C\log\dfrac{X}{a}, & n = 1 \\ \dfrac{1}{1-n}\left[\dfrac{1}{X^{n-1}} - \dfrac{1}{a^{n-1}}\right], & n \neq 1 \end{cases}$$

and therefore

$$\int_a^\infty \frac{C}{x^n}dx = \lim_{X\to\infty}\int_a^X \frac{C}{x^n}dx = \begin{cases} +\infty & n \leq 1 \\ \dfrac{C}{(n-1)a^{n-1}} & n > 1 \end{cases}$$

Thus, $\int_a^\infty \frac{C}{x^n}dx$ converges if and only if $n > 1$.

The integral is widely used as a comparison integral when testing the convergence of improper integrals.

4.4 With the help of § 11.4.2 and the comparison integral of § 11.4.3, we deduce two comparison tests of much practical utility.

I. *If f is positive in [a, x] then the integral* $\int_a^\infty f\,dx$ *converges, if there exists a positive number n greater than* 1 *and a fixed positive number M such that*

$$f(x) \le M/x^n, \text{ for every } x \ge a$$

Also the integral diverges if there exists a positive number M such that

$$f(x) \ge M/x^n, \text{ for every } x \ge a.$$

II. *If* $\lim_{x\to\infty} x^n f(x)$ *exists and is non-zero finite, then the integral* $\int_a^\infty f\,dx$ *converges if and only if* $n > 1$.

Example 11. Examine the convergence of

(*i*) $\displaystyle\int_1^\infty \frac{dx}{x\sqrt{x^2+1}}$ (*ii*) $\displaystyle\int_0^\infty \frac{x^2\,dx}{\sqrt{x^5+1}}$

(*iii*) $\displaystyle\int_0^\infty e^{-x^2}\,dx$ (*iv*) $\displaystyle\int_1^\infty \frac{\log x}{x^2}\,dx$

(*v*) $\displaystyle\int_1^\infty x^2 e^{-x}\,dx$ (*vi*) $\displaystyle\int_0^\infty \frac{\sin^2 x}{x^2}\,dx$

■ (*i*) Let $f(x) = \dfrac{1}{x\sqrt{x^2+1}}$, (behaves like x^{-2} at ∞) and $g(x) = 1/x^2$,

so that

$$\lim_{x\to\infty} \frac{f(x)}{g(x)} = \lim_{x\to\infty} \frac{x^2}{x\sqrt{x^2+1}} = 1, \text{ (non-zero finite)}$$

Hence, the two integrals $\int_1^\infty f\,dx$ and $\int_1^\infty g\,dx$ behave alike.

As $\displaystyle\int_1^\infty \frac{dx}{x^2}$ converges, therefore $\displaystyle\int_1^\infty \frac{dx}{x\sqrt{x^2+1}}$ also converges.

(*ii*) Let $f(x)=\dfrac{x^2}{\sqrt{x^5+1}}$, $\left(\sim x^{-1/2}\right)$ and $g(x)=\dfrac{1}{\sqrt{x}}$.

$$\frac{f(x)}{g(x)}=\frac{x^{5/2}}{\sqrt{x^5+1}}=\frac{1}{\sqrt{1+x^{-5}}}\to 1 \text{ as } x\to\infty$$

As $\int_0^\infty g\,dx=\int_0^\infty \frac{dx}{\sqrt{x}}$ diverges, therefore by comparison test, $\int_0^\infty \frac{x^2\,dx}{\sqrt{x^5+1}}$ also diverges.

(*iii*) 0 is not a point of infinite discontinuity and so we have to examine convergence at ∞ only.

Let us consider the integral $\int_1^\infty e^{-x^2}\,dx$.

We know

$$e^{x^2}>x^2,\ \forall \text{ real } x$$

$$\therefore \qquad e^{-x^2}<\frac{1}{x^2}$$

As $\int_1^\infty \frac{1}{x^2}\,dx$ converges at ∞, the integral $\int_1^\infty e^{-x^2}\,dx$ and therefore the integral $\int_0^\infty e^{-x^2}\,dx$ converges.

Note: $\int_0^\infty e^{-x^2}\,dx$ is called the *Euler-Poisson* integral, and it will be shown later (Double integrals) that its value is $\sqrt{\pi/2}$.

(*iv*) Here

$$x^{3/2}\,\frac{\log x}{x^2}=\frac{\log x}{x^{1/2}}\to 0 \text{ as } x\to\infty, \text{ and } \int_1^\infty \frac{dx}{x^{3/2}} \text{ converges}$$

Hence by comparison with $\int_1^\infty \frac{dx}{x^{3/2}}$, the integral $\int_1^\infty \frac{\log x}{x^2}\,dx$ also converges.

(*v*) Now

$$x^2\cdot x^n e^{-x}=x^{n+2}/e^x\to 0 \text{ as } x\to\infty, \text{ for all } n$$

and $\int_1^\infty \frac{dx}{x^2}$ converges.

Hence by comparison test, $\int_1^\infty x^n e^{-x}\,dx$ converges.

(*vi*) 0 is not a point of infinite discontinuity, so if we put

$$\int_0^\infty \frac{\sin^2 x}{x^2}\,dx = \int_0^1 \frac{\sin^2 x}{x^2}\,dx + \int_1^\infty \frac{\sin^2 x}{x^2}\,dx,$$

the first integral on the right being proper, we test the second,

$$\int_1^\infty \frac{\sin^2 x}{x^2}\,dx \text{ for convergence at } \infty.$$

Now $\frac{\sin^2 x}{x^2} \le \frac{1}{x^2}$ and $\int_1^\infty \frac{dx}{x^2}$ is convergent, therefore the integral $\int_1^\infty \frac{\sin^2 x}{x^2}\,dx$ is also convergent.

Hence, the integral $\int_0^\infty \frac{\sin^2 x}{x^2}\,dx$ is convergent.

Note: Comparison test II (limit form) cannot be used here, for $\lim f(x)/g(x)$, where $g(x) = 1/x^2$ does not exist.

Example 12. Test for convergence the integrals

(*i*) $\int_0^\infty \frac{x \tan^{-1} x}{\left(1+x^4\right)^{1/3}}\,dx$ (*ii*) $\int_{e^2}^\infty \frac{dx}{x \log \log x}$

■ (*i*) Let $f(x) = \frac{x \tan^{-1} x}{\left(1+x^4\right)^{1/3}}$

$$= \frac{\tan^{-1} x}{x^{1/3}\left(1-x^{-4}\right)^{1/3}} \left(\sim x^{-1/3} \text{ at } \infty\right)$$

and $g(x) = \frac{1}{x^{1/3}}$,

so that

$$\frac{f(x)}{g(x)} = \frac{\tan^{-1} x}{\left(1+x^{-4}\right)^{1/3}} \to \frac{\pi}{2} \text{ as } x \to \infty$$

Hence, $\int_1^\infty f\,dx$ and $\int_1^\infty g\,dx$ behave alike.

Since $\int_1^\infty \frac{dx}{x^{1/3}}$ diverges, therefore $\int_1^\infty \frac{x \tan^{-1} x\,dx}{\left(1+x^4\right)^{1/3}}$ also diverges.

(*ii*) Putting $\log x = t$ or $x = e^t$, we get

$$\int_{e^2}^{\infty} \frac{dx}{x \log \log x} = \int_{2}^{\infty} \frac{dt}{\log t}$$

which diverges by comparison with the divergent integral $\int_k^{\infty} \frac{dt}{t^m},\ m \le 1$.

Example 13. *Gamma function.* The integral $\int_0^{\infty} x^{m-1} e^{-x}\, dx$ is convergent if and only if $m > 0$.

■ Let $f(x) = x^{m-1} e^{-x} = \dfrac{e^{-x}}{x^{1-m}}$.

The integrand f has infinite discontinuity at 0 if $m < 1$. So we have to examine convergence at 0 and ∞ both.

Putting

$$\int_0^{\infty} x^{m-1} e^{-x}\, dx = \int_0^{1} x^{m-1} e^{-x}\, dx + \int_1^{\infty} x^{m-1} e^{-x}\, dx,$$

we test the two integrals on the right for convergence at 0 and ∞ respectively.

Convergence at 0, $m < 1$.

Let $g(x) = \dfrac{1}{x^{1-m}}$ so that

$$\frac{f(x)}{g(x)} = e^{-x} \to 1 \text{ as } x \to 0$$

Also $\int_0^1 g\, dx = \int_0^1 \frac{dx}{x^{1-m}}$, converges if and only if $1 - m < 1$, *i.e.*, $m > 0$.

Hence $\int_0^1 x^{m-1} e^{-x}\, dx$ converges if and only if $m > 0$.

Convergence at ∞ (ref. Example 11.11).

Let $g(x) = \dfrac{1}{x^2}$, so that

$$\frac{f(x)}{g(x)} = \frac{x^{m+1}}{e^x} \to 0 \text{ as } x \to \infty, \text{ for all } m$$

As $\int_1^\infty \frac{dx}{x^2}$ converges, therefore $\int_1^\infty x^{m-1}e^{-x}$ also converges for all m.

Hence $\int_0^\infty x^{m-1}e^{-x}\,dx$ is convergent if and only if $m > 0$.

This integral $\int_0^\infty x^{m-1}e^{-x}\,dx,\ m > 0$ called *Gamma function*, is denoted by $\Gamma(m)$.

Example 14. Examine for convergence

$$\int_0^\infty \left(\frac{1}{x} - \frac{1}{\sinh x}\right)\frac{dx}{x}$$

■ Let $f(x) = \left(\frac{1}{x} - \frac{1}{\sinh x}\right)\Big/ x$

Using L'Hospital rules, we find that $f(x) \to \frac{1}{6}$ as $x \to 0$.

Therefore 0 is not a point of infinite discontinuity of f.

To examine the convergence at ∞, we put

$$f(x) = \left(\frac{1}{x} - \frac{1}{\sinh x}\right)\Big/ x = \frac{1}{x^2} - \frac{2e^{-x}}{x\left(1-e^{-2x}\right)}$$

so that $f(x)$ behaves as $\frac{1}{x^2}$ at ∞.

By comparison with the convergent integral $\int_1^\infty \frac{dx}{x^2}$ we can easily show that $\int_1^\infty f\,dx$ and therefore $\int_0^\infty f\,dx$ is also convergent.

Example 15. Test for convergence the integral

$$\int_0^1 x^p\left(\log\frac{1}{x}\right)^q dx$$

■ Making the substitution $\log\frac{1}{x} = t$ or $x = e^{-t}$, we get

$$\int_0^1 x^p\left(\log\frac{1}{x}\right)^q dx = \int_0^\infty t^q\, e^{-(p+1)t}\,dt$$

The integrand of the last integral has a point of infinite discontinuity 0. So, we put

$$\int_0^\infty t^q e^{-(p+1)t}\,dt = \int_0^1 t^q e^{-(p+1)t}\,dt + \int_1^\infty t^q e^{-(p+1)t}\,dt$$

and examine the integrals on the right for convergence at 0 and ∞ respectively.

Convergence at 0.

Let $f(t) = t^q e^{-(p+1)t} = \dfrac{e^{-(p+1)t}}{t^{-q}}$.

By comparison with $\displaystyle\int_0^1 \frac{dt}{t^{-q}}$, we find that $\displaystyle\int_0^1 t^q e^{-(p+1)t}\,dt$ converges at 0 only if $-q < 1$ or $q > -1$, for all values of p.

Convergence at ∞.

Let $f(t) = t^q e^{-(p+1)t} = \dfrac{t^q}{e^{(p+1)t}}$ and $g(t) = \dfrac{1}{t^2}$, so that

$$\frac{f(t)}{g(t)} = \frac{t^{q+2}}{e^{(p+1)t}} \to 0 \text{ as } t \to \infty, \text{ for all } q \text{ and } p+1>0$$

Also $\displaystyle\int_1^\infty \frac{dt}{t^2}$ converges.

Hence $\displaystyle\int_1^\infty t^q e^{-(p+1)t}\,dt$ converges for all q and only if $p+1>0$.

Hence $\displaystyle\int_1^\infty t^q e^{-(p+1)t}\,dt$ and therefore $\displaystyle\int_0^1 x^p \left(\log\frac{1}{x}\right)^q dx$ is convergent for $p>-1$, $q>-1$, and divergent for all other values of p and q.

4.5 Extending methods of evaluating proper integrals to the case of improper integrals. When evaluating improper integrals we can change variables, integrate by parts, etc., *i.e.*, apply all the methods of evaluating proper integrals, provided that all the integrals entering into them are convergent.

Let us consider the integral $\displaystyle\int_3^\infty \frac{dx}{x^2+x-2}$.

The integral is convergent as can be seen by comparison with $\displaystyle\int_3^\infty \frac{dx}{x^2}$.

Again, let us decompose the integrand into partial fraction

$$\frac{1}{x^2+x-2}=\frac{1}{3(x-1)}-\frac{1}{3(x+2)} \qquad ...(1)$$

It is obvious that the integrals $\int_3^\infty \frac{dx}{x+2}$ and $\int_3^\infty \frac{dx}{x-1}$ are divergent.

Therefore, the equality

$$\int_3^\infty \frac{dx}{x^2+x-2}=\frac{1}{3}\int_3^\infty \frac{dx}{x-1}-\frac{1}{3}\int_3^\infty \frac{dx}{x+2}$$

is incorrect.

To use decomposition (1) for evaluating the integral, we can integrate (1) from 0 to X and then transform the R.H.S. of the resulting equality to the form

$$\int_3^X \frac{dx}{x^2+x-2}=\frac{1}{3}\int_3^X \frac{dx}{x-1}-\frac{1}{3}\int_3^X \frac{dx}{x+2}$$

$$=\frac{1}{2}\log\left[\frac{5X-1}{2X+2}\right]$$

Proceeding to limits as $X \to \infty$, we get

$$\int_3^\infty \frac{dx}{x^2+x-2}=\frac{1}{3}\log\frac{5}{2}$$

Example 16. Discuss the convergence of $\int_0^1 \log\overline{|x}\, dx$ and hence evaluate it.

■ $x=0$ is the only singular point of the integral.

Now

$$\Gamma(x)=\int_0^\infty t^{x-1}e^{-t}\,dt,\ x>0, \text{ thus } \overline{|x+1}=x\overline{|x},\ \ \forall\ x>0.$$

$$\therefore \qquad \log\overline{|x+1}=\log\left(x\overline{|x}\right)=\log x+\log\overline{|x}$$

or $$\int_0^1 \log\overline{|x}\,dx=\int_0^1 \log\overline{|x+1}\,dx-\int_0^1 \log x\,dx=I_1-I_2.$$

I_1 is a proper integral and $I_2 = \int_0^1 \log x\, dx$ is convergent

$$\left(\text{since } -\int_0^1 \log x\, dx = \int_0^1 \log \frac{1}{x}\, dx < \int_0^1 \frac{\varepsilon}{x^{1/2}}\, dx \left[\because \lim_{x\to 0} x^{1/2} \log\frac{1}{x} = 0\right] \text{ is convergent.}\right)$$

Hence, $\int_0^1 \log\Gamma(x)\, dx$ is convergent.

Again

$$I = \int_0^1 \log\Gamma(x)\, dx$$

$$I = \int_0^1 \log\Gamma(1-x)\, dx$$

$$\therefore \qquad 2I = \int_0^1 \log\Gamma(x) + \int_0^1 \log\Gamma(1-x)\, dx = \int_0^1 \log\Gamma(x)\Gamma(1-x)\, dx$$

We can prove that (see Appendix 1.)

$$\Gamma(p)\Gamma(1-p) = \frac{\pi}{\sin p\pi}, \text{ for } 0 < p < 1$$

$$\therefore \qquad 2I = \int_0^1 \log\left(\frac{\pi}{\sin x\pi}\right) dx, \text{ put } x\pi = z$$

$$= \frac{1}{\pi}\int_0^{\pi} \log\frac{\pi}{\sin z}\, dz = \frac{1}{\pi}\int_0^{\pi} \log\pi - \frac{1}{\pi}\int_0^{\pi} \log\sin z\, dz$$

$$= \log\pi - \frac{2}{\pi}\int_0^{\pi/2} \log\sin z\, dz$$

$$= \log\pi - \frac{2}{\pi}\left(-\frac{\pi}{2}\log 2\right) = \log 2\pi$$

$$\therefore \qquad I = \frac{1}{2}\log 2\pi.$$

EXERCISE

1. Test for convergence of the integrals:

(*i*) $\displaystyle\int_0^\infty \frac{x}{x^3+1}\,dx$ (*ii*) $\displaystyle\int_1^\infty \frac{x^3+1}{x^4}\,dx$

(*iii*) $\displaystyle\int_0^\infty \sqrt{xe^{-x}}\,dx$ (*iv*) $\displaystyle\int_e^\infty \frac{dx}{x(\log x)^{3/2}}$

(*v*) $\displaystyle\int_{-\infty}^\infty e^{-(x-a/x)^2}\,dx$.

2. Examine the convergence of

(*i*) $\displaystyle\int_0^\infty \frac{x^{2m}}{1+x^{2n}}\,dx,\ m, n > 0$ (*ii*) $\displaystyle\int_1^\infty t^m e^{-nt}\,dt$

(*iii*) $\displaystyle\int_0^\infty x^{2n+1} e^{-x^2}\,dx$, n is a positive integer (*iv*) $\displaystyle\int_0^\infty \frac{x^{p-1}}{1+x}\,dx$.

3. Show that the integral $\displaystyle\int_0^\infty \left(\frac{1}{e^x-1} - \frac{1}{x} + \frac{1}{2}\right)\frac{e^{-kx}}{x}\,dx$ converges.

4. Prove that $\displaystyle\int_0^\infty \frac{x\log x}{(1+x^2)^2}\,dx$ converges to 0.

[*Hint*: For evaluating, express as the integral $\displaystyle\int_0^\infty$ as $\displaystyle\int_0^1 + \int_1^\infty$ and put $x = 1/t$ in the second term.]

5. Show that the improper integral $\displaystyle\int_0^\infty \log(1+2\operatorname{sech} x)\,dx$ is convergent.

6. Show that $\displaystyle\int_0^\infty \left(\frac{1}{1+x} - e^{-x}\right)\frac{dx}{x}$ is convergent.

7. Putting $\log\log x = t$, in the integral $\displaystyle\int_e^\infty \frac{dx}{x^p(\log x)^q(\log\log x)^r}$, show that the integral converges for $p > 1$ (any q) only if $r < 1$. If $p = 1$ it converges only if $r < 1$ and $q > 1$. But if $p < 1$, the integral diverges for any r and q.

ANSWERS

1. (*i*), (*iii*), (*iv*), (*v*) converge, (*ii*) diverges.

2. All converge for (*i*) $n - m > \frac{1}{2}$ (*ii*) $n > 0$ (*iii*) all n (*iv*) $0 < p < 1$.

4.6 General Test for Convergence at ∞ (*Integrand may change sign*)

Cauchy's Test. *The integral* $\int_a^{\infty} f\,dx$ *converges at* ∞ *if and only if for every* $\varepsilon > 0$ *there corresponds a positive number* X_0 *such that*

$$\left| \int_{X_1}^{X_2} f dx \right| < \varepsilon, \text{ for all } X_1, X_2 > X_0$$

[It implies that $\int_{X_1}^{X_2} f\,dx \to 0$ as $X_1, X_2 \to \infty$.]

The improper integral $\int_a^{\infty} f\,dx$ exists if $\lim_{X\to\infty} \int_a^{x} f\,dx$ exists finitely.

Let $F(X) = \int_a^{x} f\,dx$, a function of X.

According to Cauchy's criterion for finite limits (§ 1.3 Ch. 5) $F(X)$ tends to a finite limit as $X \to \infty$ if and only if for every $\varepsilon > 0$ there corresponds X_0 such that for all $X_1, X_2 > X_0$,

$$\left| F(X_1) - F(X_2) \right| < \varepsilon$$

i.e.,

$$\left| \int_a^{X_1} f\,dx - \int_a^{X_2} f\,dx \right| < \varepsilon$$

or

$$\left| \int_{X_1}^{X_2} f\,dx \right| < \varepsilon.$$

Example 17. Show that $\int_0^{\infty} \frac{\sin x}{x}\,dx$ is convergent.

■ 0 is not a point of infinite discontinuity, for $\frac{\sin x}{x} \to 1$ as $x \to 0$. So, let us put

$$\int_0^{\infty} \frac{\sin x}{x}\,dx = \int_0^{1} \frac{\sin x}{x}\,dx + \int_1^{\infty} \frac{\sin x}{x}\,dx$$

and test $\int_0^{\infty} \frac{\sin x}{x}\, dx$ for convergence at ∞.

For any $\varepsilon > 0$, let X_1, X_2 be two numbers, both greater than $2/\varepsilon$.

Now

$$\int_{X_1}^{X_2} \frac{\sin x}{x}\, dx = \left[-\frac{\cos x}{x}\right]_{X_1}^{X_2} - \int_{X_1}^{X_2} \frac{\cos x}{x^2}\, dx$$

so that

$$\left|\int_{X_1}^{X_2} \frac{\sin x}{x} dx\right| \le \left|\frac{\cos X_1}{X_1} - \frac{\cos X_2}{X_2}\right| + \left|\int_{X_1}^{X_2} \frac{\cos x}{x^2} dx\right|$$

$$\le \frac{1}{X_1} + \frac{1}{X_2} + \int_{X_1}^{X_2} \frac{dx}{x^2}$$

$$< 2 \cdot \frac{\varepsilon}{2} = \varepsilon$$

Thus by Cauchy's Test $\int_1^{\infty} \frac{\sin x}{x}\, dx$ and consequently $\int_0^{\infty} \frac{\sin x}{x}\, dx$ is convergent.

4.7 Absolute Convergence

Definition. *The improper integral* $\int_a^{\infty} f\, dx$ *is said to be absolutely convergent if* $\int_a^{\infty} |f|\, dx$ *is convergent.*

With the help of Cauchy's Test, we can easily prove the following theorem.

Theorem 4. $\int_a^{\infty} f\, dx$ *exists, if* $\int_a^{\infty} |f|\, dx$ *exists.*

The converse, however, is not true. Integrals which are convergent but not absolutely are called *conditionally convergent* integrals.

Example 18. Show that $\int_1^{\infty} \frac{\sin x}{x^p} dx$ converges absolutely if $p > 1$.

■ Now

$$\left|\frac{\sin x}{x^p}\right| = \frac{|\sin x|}{x^p} \le \frac{1}{x^p} \quad \forall\, x \ge 1$$

Also $\int_1^\infty \frac{dx}{x^p}$ converges only if $p > 1$.

Therefore $\int_1^\infty \left|\frac{\sin x}{x^p}\right| dx$ converges if $p > 1$.

Hence $\int_1^\infty \frac{\sin x}{x^p}\, dx$ converges absolutely if $p > 1$.

For $p = 1$ it is not absolutely convergent.

5. INTEGRAND AS A PRODUCT OF FUNCTIONS (*Convergence at* ∞)

5.1 A Test for Absolute Convergence

A function ϕ *is bounded in* $[a, \infty[$ *and integrable in* $[a, X]$ *where* X *is a number* $\geq a$. *If* $\int_a^\infty f\, dx$ *is absolutely convergent at* ∞, then $\int_a^\infty f\phi\, dx$ *is also absolutely convergent at* ∞.

Since ϕ is bounded in $[a, \infty[$, there exists a positive K such that

$$\phi(x) \leq K, \quad \forall\, x \geq a \qquad ...(1)$$

Again, since $|f|$ is positive and $\int_a^\infty |f|\, dx$ is convergent, a number M exists such that

$$\int_a^X |f|\, dx \leq M, \text{ for all } X \geq a \qquad ...(2)$$

Using equation (1), we have

$$|f\phi| \leq K|f|, \quad \forall\, x \geq a$$

$$\therefore \quad \int_a^X |f\phi|\, dx \leq K \int_a^X |f|\, dx$$

$$\leq KM, \quad \text{for all } X \geq a,$$

which implies that the positive function $\int_a^X |f\phi|\, dx$ is bounded above by KM, for $X \geq a$.

Hence $\int_a^\infty |f\phi|\, dx$ converges.

Example 19. Discuss the convergence of the integral

$$\int_1^\infty f(x)\,dx, \text{ where } f(x) = \begin{cases} \dfrac{1}{x^2}, & x \text{ is a rational number} \\ -\dfrac{1}{x^2}, & x \text{ is an irrational number} \end{cases}$$

■ Since $\int_1^\infty |f(x)|\,dx = \int_1^\infty \frac{1}{x^2}\,dx$

is convergent, and every absolutely convergent integral is convergent. Therefore the given integral is convergent.

5.2 Tests for Convergence

Abel's Test. *If ϕ is bounded and monotonic in $[a, \infty[$, and $\int_a^\infty f\,dx$ is convergent at ∞, then $\int_a^\infty f\phi\,dx$ is convergent at ∞.*

Or

An infinite integral which converges (not necessarily absolutely) will remain convergent after the insertion of a factor which is bounded and monotonic.

Since ϕ is monotonic in $[a, \infty[$, it is integrable in $[a, X]$, for all $X \geq a$. Also, since f is integrable in $[a, X]$, we have by Second Mean Value Theorem

$$\int_{X_1}^{X_2} f\phi\,dx = \phi(X_1)\int_{X_1}^{\xi} f\,dx + \phi(X_2)\int_{\xi}^{X_2} f\,dx \qquad ...(1)$$

for $a < X_1 \leq \xi \leq X_2$.

Let $\in > 0$ be arbitrary.

Since ϕ is bounded in $[a, \infty[$, a positive number K exists such that

$$|\phi(x)| \leq K, \quad \forall\, x \geq a$$

In particular,

$$|\phi(X_1)| \leq K, |\phi(X_2)| \leq K \qquad ...(2)$$

Again, since $\int_a^\infty f\,dx$ is convergent (by § 4.6), a number X_0 exists such that

$$\left|\int_{X_1}^{X_2} f\,dx\right| < \frac{\varepsilon}{2K}, \text{ for all } X_1, X_2 \geq X_0 \qquad ...(3)$$

Let the numbers X_1, X_2 in (1) be $\geq X_0$ so that the number ξ which lies between X_1 and X_2 is also $\geq X_0$.

Hence from (3),

$$\left|\int_{X_1}^{\xi} f\,dx\right| < \frac{\varepsilon}{2K}, \left|\int_{\xi}^{X_2} f\,dx\right| < \frac{\varepsilon}{2K} \qquad ...(4)$$

Thus, from equations (1), (2) and (4), we deduce that a positive number X_0 exists such that for all $X_1, X_2 \geq X_0$,

$$\left|\int_{X_1}^{X_2} f\phi\,dx\right| \leq |\phi(X_1)| \cdot \left|\int_{X_1}^{\xi} f\,dx\right| + |\phi(X_2)| \cdot \left|\int_{\xi}^{X_2} f\,dx\right|$$

$$< K\frac{\varepsilon}{2K} + K\frac{\varepsilon}{2K} = \varepsilon$$

Hence, by Cauchy's test, $\int_a^{\infty} f\phi\,dx$ is convergent at ∞.

Dirichlet's Test. *If ϕ is bounded and monotonic in $[a, \infty[$ and tends to 0 as $x \to \infty$, and $\int_a^X f\,dx$ is bounded for $X \geq a$, then $\int_a^{\infty} f\phi\,dx$ is convergent at ∞.*

Or

An infinite integral which oscillates finitely becomes convergent after the insertion of a bounded monotonic factor which tends to zero as a limit.

Since ϕ is monotonic, it is integrable in $[a, X]$ for all $X \geq a$. Also, since f is integrable in $[a, X]$,

$\therefore$ by Second Mean Value Theorem,

$$\int_{X_1}^{X_2} f\phi\,dx = \phi(X_1)\int_{X_1}^{\xi} f\,dx + \phi(X_2)\int_{\xi}^{X_2} f\,dx \qquad ...(1)$$

for $a < X_1 \leq \xi \leq X_2$.

Again, since $\int_a^X f\,dx$ is bounded when $X \geq a$, there exists a number $K > 0$ such that

$$\left|\int_a^X f\,dx\right| \leq K, \ \forall\ X \geq a$$

$$\therefore \qquad \left|\int_{X_1}^{\xi} f\,dx\right| = \left|\int_a^{\xi} f\,dx - \int_a^{X_1} f\,dx\right|$$

$$\leq \left| \int_{X_1}^{\xi} f\,dx \right| + \left| \int_{a}^{X_1} f\,dx \right|$$

$$\leq K + K = 2K, \text{ for } X_1, \xi \geq a \qquad ...(2)$$

Similarly

$$\left| \int_{\xi}^{X_2} f\,dx \right| \leq 2K, \text{ for } X_2, \xi \geq a \qquad ...(3)$$

Let $\varepsilon > 0$ be arbitrary.

Since $\phi \to 0$ as $x \to \infty$, a positive number X_0 exists such that

$$|\phi(X_1)| < \frac{\varepsilon}{4K}, |\phi(X_2)| < \frac{\varepsilon}{4K} \text{ where } X_2 \geq X_1 \geq X_0 \qquad ...(4)$$

Let the numbers X_1, X_2 in (1) be $\geq X_0$, so that from (1), (2), (3) and (4), we get

$$\left| \int_{X_1}^{X_2} f\phi\,dx \right| \leq |\phi(X_1)| \left| \int_{X_1}^{\xi} f\,dx \right| + |\phi(X_2)| \left| \int_{\xi}^{X_2} f\,dx \right|$$

$$< \frac{\varepsilon}{4K} 2K + \frac{\varepsilon}{4K} 2K = \varepsilon$$

Hence, by Cauchy's test, $\int_a^{\infty} f\phi\,dx$ is convergent at ∞.

ILLUSTRATIONS

1. The integral $\int_1^{\infty} \frac{\sin x}{x^p}\,dx$ is convergent for $p > 0$.

   Let $f(x) = \sin x$, and $\phi(x) = \frac{1}{x^p}$.

   Here

$$\left| \int_1^X f\,dx \right| = \left| \int_1^X \sin x\,dx \right| = |\cos 1 - \cos X|$$

$$\leq |\cos 1| + |\cos X| \leq 2, \text{ for } 1 \leq X < \infty.$$

   Also $\phi(x) = \frac{1}{x^p}$ is a monotone decreasing function tending to 0 as $x \to \infty$ for $p > 0$.

Therefore, by Dirichlet's test, $\int_1^\infty f\phi\,dx = \int_1^\infty \frac{\sin x}{x^p}\,dx$ converges for $p > 0$.

Earlier, in Example 18, it was shown that $\int_1^\infty \frac{\sin x}{x^p}\,dx$ converges absolutely for $p > 1$.

Thus we conclude that $\int_1^\infty \frac{\sin x}{x^p}\,dx$ converges absolutely for $p > 1$, but only conditionally for $0 < p \le 1$.

2. The integral $\int_1^\infty \frac{\sin x \log x}{x}\,dx$ is convergent.

Let $f(x) = \sin x$, $\phi(x) = \frac{\log x}{x}$.

Now $\left|\int_e^\infty \sin x\,dx\right|$ is bounded above by 2, and ϕ is monotone decreasing to 0 as $x \to \infty$.

Hence, the given integral converges by Dirichlet's test.

EXERCISE

1. Establish the convergence of the integrals

(*i*) $\int_0^\infty e^{-px}\frac{\sin x}{x}\,dx,\ p \ge 0$ (*ii*) $\int_1^\infty \left(1 - e^{-x}\right)\frac{\cos x}{x}\,dx$

2. Show the convergence of

(*i*) $\int_0^\infty \frac{\sin kx}{x}\,dx$ (*ii*) $\int_0^\infty \frac{\sin x}{\sqrt{x}}\,dx$

(*iii*) $\int_0^\infty e^{-a^2x^2}\cos bx\,dx$ (*iv*) $\int_0^\infty e^{-a^2x^2}\sin 2bx\frac{dx}{x}$

(*v*) $\int_0^\infty \frac{\cos x}{\sqrt{x^2 + x}}\,dx$ (*vi*) $\int_0^\infty \cos x^2\,dx$

Example 20. ***Second method.*** Show that $\int_0^\infty \frac{\sin x}{x}\,dx$ is convergent, but not absolutely.

- Here, 0 is not a point of infinite discontinuity because $\frac{\sin x}{x} \to 1$ as $x \to 0$. So, let us put

$$\int_0^\infty \frac{\sin x}{x}\,dx = \int_0^1 \frac{\sin x}{x}\,dx + \int_1^\infty \frac{\sin x}{x}\,dx$$

Now $\int_0^1 \frac{\sin x}{x}\,dx$ is a proper integral.

To examine the convergence of $\int_1^\infty \frac{\sin x}{x}\,dx$ at ∞, we see that

$$\left|\int_1^X \sin x\,dx\right| = |\cos 1 - \cos X| \le |\cos 1| + |\cos X| < 2$$

so that $\left|\int_1^X \sin x\,dx\right|$ is bounded above for all $X \ge 1$.

Also, $1/x$ is a monotone decreasing function tending to 0 as $x \to \infty$.

Hence by Dirichlet's test, $\int_1^\infty \frac{\sin x}{x}\,dx$ is convergent.

Hence $\int_0^\infty \frac{\sin x}{x}\,dx$ is convergent.

To show that $\int_0^\infty \frac{\sin x}{x}\,dx$ is not absolutely convergent, we proceed as follows:

Consider for $n \ge 1$, the proper integral

$$\int_0^{n\pi} \left|\frac{\sin x}{x}\right| dx = \sum_{r=1}^{n} \int_{(r-1)\pi}^{r\pi} \frac{|\sin x|}{x}\,dx$$

Now $\forall\, x \in [(r-1)\pi, r\pi]$

$$\int_{(r-1)\pi}^{r\pi} \frac{|\sin x|}{x}\,dx \ge \int_{(r-1)\pi}^{r\pi} \frac{|\sin x|}{r\pi}\,dx$$

Putting $x = (r-1)\pi + y$

$$\int_{(r-1)\pi}^{r\pi} \frac{|\sin x|\,dx}{r\pi} = \int_0^{\pi} \frac{|\sin((r-1)\pi + y)|\,dy}{r\pi}$$

$$= \frac{1}{r\pi}\int_0^{\pi} \sin y\, dy = \frac{2}{r\pi}$$

Hence,

$$\int_0^{n\pi} \frac{|\sin x|}{x}\, dx = \sum_{r=1}^{n} \int_{(r-1)\pi}^{r\pi} \frac{|\sin x|\, dx}{x} \geq \sum_{r=1}^{n} \frac{2}{r\pi}$$

But $\sum_{r=1}^{n} \frac{2}{r\pi}$ is a divergent series.

$\therefore$
$$\lim_{n\to\infty} \int_0^{n\pi} \frac{|\sin x|}{x}\, dx \geq \lim_{n\to\infty} \sum_{r=1}^{n} \frac{2}{r\pi}$$

$\Rightarrow$
$$\lim_{n\to\infty} \int_0^{n\pi} \frac{|\sin x|}{x}\, dx \text{ is infinite.}$$

Now, let t be a real number. There exist positive integer n such that

$$n\pi \leq t < (n+1)\pi$$

We have

$$\int_0^{t} \frac{|\sin x|}{x}\, dx \geq \int_0^{n\pi} \frac{|\sin x|}{x}\, dx$$

Let $t \to \infty$, so that $n \to \infty$, thus we see that

$$\int_0^{t} \frac{|\sin x|}{x}\, dx \to \infty$$

$\Rightarrow$
$$\int_0^{\infty} \frac{|\sin x|}{x}\, dx \text{ does not converge.}$$

Ex. Show that $\int_0^{\infty} \frac{\sin x}{x^p} dx,\ 0 < p \leq 1$ is convergent, but not absolutely.

Example 21. Test the convergence of $\int_0^{\infty} \frac{\sin x^m}{x^n}\, dx$.

■ For $m = 0$, the integral reduces to $\int_0^{\infty} \frac{dx}{x^n}$, which converges at 0 for $n < 1$, and converges at ∞ for $n > 1$.

Thus the integral cannot converge at 0 and ∞ both. Hence it diverges for $m = 0$.

Let $m \neq 0$.

Substituting $x^m = t$, we get

$$\int_0^\infty \frac{\sin x^m}{x^n}\,dx = \frac{1}{m}\int_0^\infty \frac{\sin t}{t^{\{(n-1)/m\}+1}}\,dt$$

- Let $f(t) = \dfrac{\sin t}{t^{\{(n-1)/m\}+1}} = \dfrac{\sin t}{t}\cdot\dfrac{1}{t^{(n-1)/m}}$.

Taking a number, say 1, greater than 0, we write

$$\int_0^\infty f\,dt = \int_0^1 f\,dt + \int_1^\infty f\,dt$$

and examine the two integrals on the right for convergence at 0 and ∞, respectively.

Convergence at 0

$\int_0^1 f\,dt$ is a proper integral for $\dfrac{n-1}{m} \le 0$ but has infinite discontinuity at 0 for $\dfrac{n-1}{m} > 0$.

For $\dfrac{n-1}{m} > 0$, let $g(t) = \dfrac{1}{t^\alpha}$.

Now

$$\lim_{t\to 0}\frac{f(t)}{g(t)} = \lim_{t\to 0}\frac{\sin t}{t}\cdot\frac{1}{t^{\{(n-1)/m\}-\alpha}}$$

$$= \text{a finite quantity, if } \frac{n-1}{m} = \alpha \le 0$$

and $\quad \int_0^1 g\,dt = \int_0^1 \dfrac{dt}{t^\alpha}$ converges if and only if $\alpha < 1$.

Hence $\int_0^1 f\,dt$ converges if $\dfrac{n-1}{m} \le \alpha < 1$.

Thus $\int_0^1 \dfrac{\sin t}{t^{\{(n-1)/m\}+1}}\,dt$ converges at 0 if $\dfrac{n-1}{m} < 1$, which includes the case $\dfrac{n-1}{m} \le 0$ when the integral is proper.

Convergence at ∞

$$\left|\int_0^X \sin t\,dt\right| \le 2, \text{ bounded}$$

and $1/t^{\{(n-1)/m\}+1}$ is a monotone decreasing function, tending to 0 as $t \to \infty$.

when $\dfrac{n-1}{m}+1>0$ or $\dfrac{n-1}{m}>-1$.

So by Dirichlet's test, $\int_1^\infty f\,dt$ converges at ∞ if $-1<\dfrac{n-1}{m}$.

Hence, the given integral converges when $-1<\dfrac{n-1}{m}<1$ [or equivalently $-1<\dfrac{n-1}{-m}<1$], *i.e.*, for all m (positive or negative), and $1-m<n<1+m$.

Ex. Discuss the absolute convergence of $\displaystyle\int_0^\infty \frac{\sin x^m}{x^n}\,dx$.

Example 22. Using $\displaystyle\int_0^\infty \frac{\sin x}{x}\,dx=\frac{1}{2}\pi$, show that

$$\int_0^\infty \frac{\sin^2 x}{x^2}\,dx=\frac{1}{2}\pi$$

■ The integral is convergent (Example 11).
To compute it, let us integrate by parts.

$$\therefore \qquad \int_0^\infty \frac{\sin^2 x}{x^2}\,dx=\left[-\frac{\sin^2 x}{x}\right]_0^\infty+\int_0^\infty \frac{\sin 2x}{x}\,dx$$

$$=\int_0^\infty \frac{\sin t}{t}\,dt=\frac{1}{2}\pi$$

Example 23. Show that $\displaystyle\int_2^\infty \frac{\cos x}{\log x}\,dx$ is conditionally convergent.

■ Let $\phi(x)=\dfrac{1}{\log x}$, $f(x)=\cos x$

$$\left|\int_2^X \cos x\,dx\right|=\left|\sin X-\sin 2\right|\le\left|\sin X\right|+\left|\sin 2\right|\le 2$$

So that $\displaystyle\int_2^X \cos x\,dx$ is bounded for all $X\ge 2$

Also $\phi(x) = \dfrac{1}{\log x}$ is a monotonic decreasing function tending to 0 as $x \to \infty$.

Hence by Dirichlet's test $\displaystyle\int_2^\infty \frac{\cos x}{\log x}\,dx$ is convergent

For absolute convergence consider

$$I = \int_2^\infty \left|\frac{\cos x}{\log x}\right| dx = \int_2^{3\pi/2} \frac{|\cos x|}{\log x} dx + \int_{3\pi/2}^{5\pi/2} \frac{|\cos x|}{\log x} dx + \dots + \int_{(2n-1)\pi/2}^{(2n+1)\pi/2} \frac{|\cos x|}{\log x} dx + \dots$$

$$\therefore \qquad I = \int_{\pi/2}^{2} \frac{|\cos x|}{\log x} dx + \int_2^{3\pi/2} \frac{|\cos x|}{\log x} dx + \dots + \int_{(2n-1)\pi/2}^{(2n+1)\pi/2} \frac{|\cos x|}{\log x} dx + \dots - \int_{\pi/2}^{2} \frac{|\cos x|}{\log x} dx$$

$$= \sum_{r=1}^{\infty} \int_{(2r-1)\pi/2}^{(2r+1)\pi/2} \frac{|\cos x|}{\log x} dx - \int_{\pi/2}^{2} \frac{|\cos x|}{\log x}$$

Now

$$\int_{(2r-1)\pi/2}^{(2r+1)\pi/2} \frac{|\cos x|}{\log x} dx \ge \frac{1}{\log(2r+1)\pi/2} \left| \int_{(2r-1)\pi/2}^{(2r+1)\pi/2} \cos x\, dx \right|$$

$$= \frac{1}{\log(2r+1)\pi/2} \left| \sin(2r+1)\pi/2 - \sin(2r-1)\pi/2 \right|$$

$$= \frac{\left|2(-1)^r\right|}{\log(2r+1)\pi/2}$$

$$= \frac{2}{\log(2r+1)\pi/2}$$

$$\therefore \qquad I \ge \sum_{r=1}^{\infty} \frac{2}{\log(2r+1)\pi/2} - \int_{\pi/2}^{2} \frac{|\cos x|}{\log x} dx$$

But $\displaystyle\sum_{x=2}^{\infty} \frac{1}{\log x}$ is divergent and $\displaystyle\int_{\pi/2}^{2} \frac{|\cos x|}{\log x}\, dx$ is a proper integral.

Hence

$$I = \int_2^\infty \frac{|\cos x|}{\log x} dx \text{ is divergent}$$

And so

$$\int_2^\infty \frac{\cos x}{\log x}\,dx \text{ is conditionally convergent.}$$

Example 24. The function f is defined on $[0, \infty[$ by

$$f(x) = (-1)^{n-1},\ n-1 \le x < n,\ n \in \mathbf{N}$$

Show that the integral $\int_0^\infty f(x)dx$ does not converge.

■ Now

$$\int_0^{2n} f(x)dx = \int_0^1 (-1)^0\,dx + \int_1^2 (-1)dx + \int_2^3 (-1)^2\,dx + \ldots + \int_{2n-1}^{2n} (-1)^{2n-1}\,dx$$
$$= 1 - 1 + 1 - 1 + 1 - \ldots + 1 - 1 = 0$$

and

$$\int_0^{2n+1} f(x)dx = \int_0^1 dx + \int_1^2 (-1)dx + \ldots \int_{2n}^{2n+1} (-1)^{2n}\,dx$$
$$= 1 - 1 + 1 \ldots - 1 + 1 = 1$$

$$\lim_{n\to\infty} \int_0^{2n} f(x)dx = 0 \text{ and } \lim_{n\to\infty} \int_0^{2n+1} f(x)dx = 1$$

Hence, the integral does not exist and therefore it is not convergent.

Example 25. The function f is defined on $]0, 1]$ by

$$f(x) = (-1)^{n+1} n(n+1), \frac{1}{n+1} \le x < \frac{1}{n},\ n \in \mathbf{N}$$

Show that $\int_0^1 f(x)dx$ does not converge.

■ $$\int_{1/2n+1}^{1} f(x)dx = \int_{1/2n+1}^{a/2n} f\,dx + \int_{1/2n}^{1/2n-1} f\,dx + \ldots + \int_{1/3}^{1/2} f\,dx + \int_{1/2}^{1} f\,dx$$

$$= \int_{1/2n+1}^{1/2n} (-1)^{2n+1} 2n(2n+1)dx + \int_{1/2n}^{1/2n-1} (-1)^{2n} (2n-1)2n\,dx + \ldots + \int_{1/3}^{1/2} (-1)^3\, 2.3\,dx + \int_{1/2}^{1} (-1)^2\, 1.2\,dx$$

$$= -1 + 1 - 1 + 1 - \ldots + 1 = 0$$

And

$$\int_{1/2n}^{1} f\,dx = \int_{1/2n}^{1/2n-1} (-1)^{2n}(2n-1)2n\,dx + \int_{1/2n-1}^{1/2n-2} (-1)^{2n-1}(2n-2)(2n-1)dx + \dots + \int_{1/2}^{1} (-1)^2 1\cdot 2\,dx$$

$$= 1 - 1 + 1 - 1 + \dots + 1 = 1$$

$$\therefore \qquad \lim_{n\to\infty} \int_{1/2n}^{1} f(x)dx = 1$$

$$\text{and} \qquad \lim_{n\to\infty} \int_{1/2n+1}^{1} f(x)\,dx = 0$$

Hence, the integral $\int_0^1 f(x)dx$ does not converge.

Example 26. Test the convergence of

(*i*) $\displaystyle\int_0^\infty \frac{xdx}{1+x^4\cos^2 x}$

(*ii*) $\displaystyle\int_0^\infty \frac{dx}{1+x^4\cos^2 x}$

■ (*i*) The integrand is positive for positive values of x but the tests obtained for the convergence of positive integrands so far, are not applicable. In order to show the integral convergent we proceed as follows:

Consider $\displaystyle\int_0^{n\pi} \frac{x\,dx}{1+x^4\cos^2 x}$ and write

$$\therefore \qquad \int_0^{n\pi} \frac{x\,dx}{1+x^4\cos^2 x} = \sum_{r=1}^{n} \int_{(r-1)\pi}^{r\pi} \frac{x\,dx}{1+x^4\cos^2 x}$$

Now $\forall\ x \in [(r-1)\pi, r\pi]$, we have

$$\frac{x}{1+x^4\cos^2 x} \geq \frac{(r-1)\pi}{1+r^4\pi^4\cos^2 x}$$

$$\therefore \qquad \int_{(r-1)\pi}^{r\pi} \frac{x\,dx}{1+x^4\cos^2 x} \geq \int_{(r-1)\pi}^{r\pi} \frac{(r-1)\pi\,dx}{1+x^4\pi^4\cos^2 x}$$

Putting $x=(r-1)\pi+y$, we see that

$$\int_{(r-1)\pi}^{r\pi}\frac{(r-1)\pi\,dx}{1+r^4\pi^4\cos^2 x}=\int_0^{\pi}\frac{(r-1)\pi\,dy}{1+r^4\pi^4\cos^2\{(r-1)\pi+y\}}$$

$$\int_0^{\pi}\frac{(r-1)\pi\,dy}{1+r^4\pi^4\cos^2 y}$$

$$=2(r-1)\pi\int_0^{\pi/2}\frac{dy}{1+r^4\pi^4\cos^2 y}$$

$$=2(r-1)\pi\int_0^{\pi/2}\frac{\sec^2 y\,dy}{1+\tan^2 y+r^4\pi^4}$$

$$=\frac{2(r-1)\pi}{\sqrt{1+r^4\pi^4}}\tan^{-1}\left(\frac{\tan y}{\sqrt{1+r^4\pi^4}}\right)\Bigg|_0^{\pi/2}=\frac{(r-1)\pi^2}{\sqrt{1+r^4\pi^4}}$$

$\therefore$
$$\sum_{r=1}^{n}\int_{(r-1)\pi}^{r\pi}\frac{x\,dx}{1+x^4\cos^2 x}\geq\sum_{r=1}^{n}\frac{(r-1)\pi^2}{\sqrt{1+r^4\pi^4}}$$

Hence,

$$\lim_{n\to\infty}\int_0^{n\pi}\frac{x\,dx}{1+x^4\cos^2 x}\geq\lim_{n\to\infty}\sum_{r=1}^{n}\frac{(r-1)\pi^2}{\sqrt{1+r^4\pi^4}}$$

But $\displaystyle\sum_{r=1}^{\infty}\frac{(r-1)\pi^2}{\sqrt{1+r^4\pi^4}}$ is a divergent series $\left(\sim\displaystyle\sum_{r=1}^{\infty}\frac{1}{r}\right)$

$\therefore$ $\displaystyle\int_0^{\infty}\frac{x\,dx}{1+x^4\cos^2 x}$ is divergent.

(ii) $\displaystyle\int_0^{\infty}\frac{dx}{1+x^4\cos^2 x}$

Consider

$$\int_0^{n\pi}\frac{dx}{1+x^4\cos^2 x}=\sum_{r=1}^{n}\int_{(r-1)\pi}^{r\pi}\frac{dx}{1+x^4\cos^2 x}$$

Now $\forall\ x \in [(r-1)\pi, r\pi]$

$$\frac{1}{1+x^4\cos^2 x} \le \frac{1}{1+(r-1)^4\pi^4\cos^2 x}$$

$$\therefore \qquad \int_{(r-1)\pi}^{r\pi} \frac{dx}{1+x^2\cos^2 x} \le \int_0^{\pi} \frac{dy}{1+(r-1)^4\pi^4\cos^2 y}$$

where $\qquad x = (r-1)\pi + y$

$$= 2\int_0^{\pi/2} \frac{\sec^2 y\,dy}{1+\tan^2 y+(r-1)^4\pi^4}$$

$$= \frac{2}{\sqrt{1+(r-1)^4\pi^4}} \tan^{-1} \frac{\tan y}{\sqrt{1+(r-1)^4\pi^4}}\Bigg|_0^{\pi/2}$$

$$= \frac{\pi}{\sqrt{1+(r-1)^4\pi^4}}$$

Hence

$$\lim_{n\to\infty} \int_0^{n\pi} \frac{dx}{1+x^4\cos^2 x} \le \lim_{n\to\infty} \sum_{r=1}^{n} \frac{\pi}{\sqrt{1+(r-1)^4\pi^4}}$$

But $\sum\limits_{r=1}^{\infty} \dfrac{\pi}{\sqrt{1+(r-1)^4\pi^4}}$ is a convergent series $\left(\sim \sum\limits_{r=1}^{\infty} \dfrac{1}{r^2}\right)$

$\therefore \displaystyle\int_0^{\infty} \frac{dx}{1+x^4\cos^2 x}$ is convergent.

EXERCISE

1. Discuss the convergence or divergence of the following integrals:

(i) $\displaystyle\int_0^{\infty} \frac{\cos\alpha x\cos\beta x}{x}\,dx$ (ii) $\displaystyle\int_0^{\infty} \frac{x^m\cos ax}{1+x^n}\,dx$

(iii) $\displaystyle\int_0^{\infty} \frac{x^p\sin^2 x}{1+x^2}\,dx$ (iv) $\displaystyle\int_1^{\infty} \sin x^p\,dx$

(v) $\displaystyle\int_0^{\infty} \frac{\sin(x+x^2)}{x^n}\,dx$ (vi) $\displaystyle\int_2^{\infty} \frac{\sin x}{\log x}\,dx$

2. Using Poisson's integral $\int_0^\infty e^{-x^2}\,dx = \frac{\sqrt{x}}{2}$ and Dirichlet's integral $\int_0^\infty \frac{\sin x}{x}\,dx = \frac{\pi}{2}$, show that

(*i*) $\int_0^\infty \frac{e^{-x}}{\sqrt{x}}\,dx = \sqrt{\pi}$

(*ii*) $\int_0^\infty x^2 e^{-x^2}\,dx = \sqrt{\pi/4}$ [Integrate by parts]

(*iii*) $\int_0^\infty \frac{\sin^3 x}{x}\,dx = \frac{\pi}{4}$ [Use: $4\sin^3 x = 3\sin x - \sin 3x$]

(*iv*) $\int_0^\infty \frac{\sin^4 x}{x^2}\,dx = \frac{\pi}{4}$.

3. Show that $\int_0^\infty \frac{x\,dx}{1 + x^6 \sin^2 x}$ converges, but $\int_0^\infty \frac{x\,dx}{1 + x^4 \sin^2 x}$ does not.

4. Show that $\int_0^\infty \frac{dx}{1 + x^4 \sin^2 x}$ converges, but $\int_0^\infty \frac{dx}{1 + x^2 \sin^2 x}$ does not.

ANSWERS

1. (*i*) Div.
   (*ii*) Conv. when $n > 0, -1 < m < n$, or $n < 0, 0 > m > n - 1$,
   (*iii*) Conv. for $-3 < p < 1$
   (*iv*) Conv. for $p > 1$
   (*v*) Conv. for $-1 < n < 2$
   (*vi*) Conv. but not absolutely.

12 Uniform Convergence

Thus far we have considered, almost exclusively, sequences and series whose terms were numbers. It was only in particularly simple cases that the terms depended on a variable. We shall now consider sequences and series, whose terms depend on a variable, *i.e.*, those whose terms are real valued functions defined on an *interval* as domain. We, accordingly, denote the terms by $f_n(x)$ and consider sequences and series of the form $\{f_n\}$ and Σf_n respectively.

1. POINTWISE CONVERGENCE

Suppose $\{f_n\}, n = 1, 2, 3, ...,$ is a sequence of functions, defined on an interval I, $a \le x \le b$. To each point $\xi \in I$, there corresponds a sequence of numbers $\{f_n(\xi)\}$ with terms

$$f_1(\xi), f_2(\xi), f_a(\xi), ...$$

Further, let us suppose that the sequence of numbers $\{f_n(\xi)\}$ converges for every $\xi \in I$.

Let $\{f_n(\xi)\}$ converges to $f(\xi)$.

This way let the sequences at (all) points $\xi, \eta, \zeta, ...,$ of I converge to

$$f(\xi), f(\eta), f(\zeta), \qquad ...(1)$$

We now define, in a natural way, a real valued function f with domain I and range the set defined by (1), so that its value $f(\eta)$ for $\eta \in I$ is $\lim \{f_n(\eta)\}$.

Thus

$$f(x) = \lim_{n\to\infty} \{f_n(x)\}, \ \forall \ x \in I \qquad ...(2)$$

The function f, so defined, is referred to as the *limit* or the *point-wise limit* of the sequence $\{f_n\}$ on $[a, b]$, and the sequence $\{f_n\}$ is said to be *pointwise convergent* to f on $[a, b]$.

Similarly, if the series Σf_n converges for every point $x \in I$, and we define

$$f(x) = \sum_{n=0}^{\infty} f_n(x), \ \forall \ x \in [a, b] \qquad ...(3)$$

the function f is called the *sum* or the *pointwise sum* of the series Σf_n on $[a, b]$.

Thus, if f is the point-wise limit of a sequence of functions $\{f_n\}$ defined on $[a, b]$, then to each $\varepsilon > 0$ and to each $x \in [a, b]$, there corresponds an integer m such that

$$|f_n(x) - f(x)| < \varepsilon, \quad \forall \ n \ge m \qquad ...(4)$$

1.1 For a sequence (series) of variable terms, the most important question will usually be whether, and to what extent, properties belonging to terms, viz., boundedness, continuity, integrability, differentiability, etc., are transferred to the limit function of the corresponding sequence (series). Let us consider a few examples.

1. The geometric series

$$1+x+x^2+x^3+\dots$$

converges to $(1-x)^{-1}$ in the interval $-1<x<1$.

All the terms are bounded without the sum being so.

2. Consider the series

$$\sum_{n=0}^{\infty} f_n, \text{ where } f_n(x)=\frac{x^2}{(1+x^2)^n} \ (x \text{ real})$$

At $x=0$, each $f_n(x)=0$, so that the sum of the series $f(0)=0$.

For $x \neq 0$, it forms a geometric series with common ratio $1/(1+x^2)$, so that its sum function $f(x)=1+x^2$.

Hence,

$$f(x)=\begin{cases}1+x^2, & x\neq 0\\ 0 \quad , & x=0\end{cases}$$

Each term of the series is continuous but the sum f is not.

3. The sequence $\{f_n\}$, where $f_n(x)=\dfrac{\sin nx}{\sqrt{n}}$ $(x$ real), has the limit

$$f(x)=\lim_{n\to\infty} f_n(x)=0$$

$\therefore$
$$f'(x)=0, \text{ and so } f'(0)=0$$

But

$$f_n'(x)=\sqrt{n}\cos nx$$

so that

$$f_n'(0)=\sqrt{n}\to\infty \quad \text{as } n\to\infty$$

Thus at $x=0$, the sequence $\{f_n'(x)\}$ diverges whereas the limit function $f'(x)=0$, *i.e.*, the limit of differentials is not equal to the differential of the limit.

4. Consider the sequence $\{f_n\}$, where

$$f_n(x)=nx(1-x^2)^n, \ \ 0\le x\le 1, n=1,2,3,\dots$$

For $0<x\le 1$, $\lim\limits_{n\to\infty} f_n(x)=0$

At $x=0$, each $f_n(0)=0$, so that $\lim\limits_{n\to\infty} f_n(0)=0$

Thus the limit function $f(x) = \lim_{n\to\infty} f_n(x) = 0$, for $0 \le x \le 1$

$\therefore$
$$\int_0^1 f(x)\, dx = 0$$

Again,
$$\int_0^1 f_n(x)\, dx = \int_0^1 nx(1 - x^2)^n\, dx = \frac{n}{2n+2}$$

so that
$$\lim_{n\to\infty} \left\{ \int_0^1 f_n(x)\, dx \right\} = \frac{1}{2}$$

Thus,
$$\lim_{n\to\infty} \left\{ \int_0^1 f_n\, dx \right\} \neq \int_0^1 f\, dx = \int_0^1 [\lim_{n\to\infty} \{f_n\}] dx.$$

Thus, the limit of integrals is not equal to the integral of the limit. In other words, the sequence of integrals may not converge to the integral of the limit of the sequence.

These few examples should convince us that a quite new category of problems arises with the consideration of sequences (series) of variable terms. We have to investigate under what supplementary conditions these or other properties of the terms f_n are transferred to the limit function f. A concept of great importance in this respect is known as *Uniform convergence* of a sequence (series) in its domain of definition, $[a, b]$.

2. UNIFORM CONVERGENCE ON AN INTERVAL

A sequence of functions $\{f_n\}$ is said to *converge uniformly* on an interval $[a, b]$ to a function f if for any $\varepsilon > 0$ and for *all* $x \in [a, b]$ there exists an integer N (independent of x but dependent on ε) such that for all $x \in [a, b]$,

$$|f_n(x) - f(x)| < \varepsilon, \quad \forall\, n \ge N \qquad ...(5)$$

It is clear that every uniformly convergent sequence is pointwise convergent, and the uniform limit function is same as the pointwise limit function.

The difference between the two concepts is this: In case of pointwise convergence, for each $\varepsilon > 0$ and for each $x \in [a, b]$ there exists an integer N (depending on ε and x both) such that (4) holds for $n \ge N$; whereas in uniform convergence, for each $\varepsilon > 0$, it is possible to find one integer N (dependent on ε alone) which will do for all $x \in [a, b]$.

Notes:

1. If a sequence converges pointwise to f then for a given $\varepsilon > 0$, each point x_i of $[a, b]$ determines an integer N_i such that

$$|f_n(x_i) - f(x_i)| < \varepsilon, \quad \text{for } n \geq N_i$$

Consideration of all points of $[a, b]$ gives rise to a sequence of integers $N_1, N_2, N_3, \ldots$

In case the sequence $\{N_i\}$ is bounded above, with supremum N, say, then (4) holds for all points of $[a, b]$ when $n \geq N$ and so the given sequence $\{f_n\}$ converges uniformly on $[a, b]$.

If no such N exists, the sequence $\{f_n\}$ is not uniformly convergent.

2. *Uniform convergence* $\Rightarrow$ *pointwise convergence*

but not vice versa. However

Non-pointwise convergence $\Rightarrow$ *non-uniform convergence*

i.e., a sequence which is not pointwise convergent cannot be uniformly convergent.

2.1 A series of functions Σf_n is said to converge uniformly on $[a, b]$ if the sequence $\{S_n\}$ of its partial sums, defined by

$$S_n(x) = \sum_{i=1}^{n} f_i(x)$$

converges uniformly on $[a, b]$.

Thus, a series of functions Σf_n *converges uniformly* to f on $[a, b]$ if for $\varepsilon > 0$ and all $x \in [a, b]$ there exists an integer N (independent of x and dependent on ε) such that for all x in $[a, b]$

$$|f_1(x) + f_2(x) + \ldots + f_n(x) - f(x)| < \varepsilon, \quad \text{for } n \geq N.$$

2.2 Cauchy's Criterion for Uniform Convergence

Theorem 1. *A sequence of functions $\{f_n\}$ defined on $[a, b]$ converges uniformly on $[a, b]$ if and only if for every $\varepsilon > 0$ and for all $x \in [a, b]$, there exists an integer N such that*

$$|f_{n+p}(x) - f_n(x)| < \varepsilon, \quad \forall\, n \geq N,\ p \geq 1 \qquad \ldots(1)$$

Necessary. Let the sequence $\{f_n\}$ uniformly converges on $[a, b]$ to the limit function f, so that for a given $\varepsilon > 0$, and for all $x \in [a, b]$, there exist integers m_1, m_2 such that

$$|f_n(x) - f(x)| < \varepsilon/2, \quad \forall\, n \geq m_1$$

and

$$|f_{n+p}(x) - f(x)| < \varepsilon/2, \quad \forall\, n \geq m_2,\ p \geq 1$$

Let $N = \max\ (m_1, m_2)$.

$$\therefore \qquad |f_{n+p}(x) - f_n(x)| \leq |f_{n+p}(x) - f(x)| + |f_n(x) - f(x)|$$

$$< \varepsilon/2 + \varepsilon/2 = \varepsilon, \quad \forall\, n \geq N,\ p \geq 1$$

Sufficient. Let now the given condition hold.

By Cauchy's general principle of convergence, $\{f_n\}$ converges for each $x \in [a, b]$ to a limit, say f. Thus the sequence converges pointwise to f. Let us now prove that the convergence is uniform.

For a given $\varepsilon > 0,$ let us choose an integer N such that (1) holds. Fix n, and let $p \to \infty$ in (1). Since $f_{n+p} \to f$ as $p \to \infty,$ we get

$$|f(x) - f_n(x)| < \varepsilon \quad \forall\, n \geq N, \text{ all } x \in [a, b]$$

which proves that $f_n(x) \to f(x)$ uniformly on $[a, b]$.

Note: In the statement of the theorem, (1) may be equivalently replaced by

$$|f_n(x) - f_m(x)| < \varepsilon, \quad \forall\, n, m \geq N$$

Theorem 2. *A series of functions Σf_n defined on $[a, b]$ converges uniformly on $[a, b]$ if and only if for every $\varepsilon > 0$ and for all $x \in [a, b],$ there exists an integer N such that*

$$|f_{n+1}(x) + f_{n+2}(x) + \ldots + f_{n+p}(x)| < \varepsilon, \quad \forall\, n \geq N,\ p \geq 1 \qquad \ldots(2)$$

The proof is left to the readers.

Note: Relation (2) in the statement may be replaced by

$$|f_{m+1}(x) + f_{m+2}(x) + \ldots + f_n(x)| < \varepsilon, \quad \forall\, n, m \geq N$$

2.3 Solved Examples

Example 1. Test for uniform convergence, the sequence $\{f_n\}$, where

$$f_n(x) = \frac{nx}{1 + n^2x^2}, \text{ for all real } x.$$

■ The sequence converges pointwise to f, where

$$f(x) = 0, \quad \forall \text{ real } x$$

Let $\{f_n\}$ converges uniformly in any interval $[a, b]$, so that the point-wise limit is also the uniform limit. Therefore for given $\varepsilon > 0,$ there exists an integer N such that for all $x \in [a, b]$

$$\left|\frac{nx}{1 + n^2x^2} - 0\right| < \varepsilon, \quad \forall\, n \geq N$$

If we take $\varepsilon = \frac{1}{3},$ and m an integer greater than N such that $1/m \in [a, b],$ we find on taking $n = m$ and $x = 1/m,$ that

$$\frac{nx}{1+n^2x^2} = \frac{1}{2} \nless \frac{1}{3} = \varepsilon$$

We, thus, arrive at a contradiction and so the sequence is not uniformly convergent in the interval $[a, b]$, which contains the point $1/m$. But since $1/m \to 0$, the interval $[a, b]$ contains 0. Hence, the sequence is not uniformly convergent on any interval $[a, b]$ containing 0.

Example 2. Show that the sequence $\{f_n\}$, where

$$f_n(x) = x^n$$

is uniformly convergent on $[0, k]$, $k < 1$ and only pointwise convergent on $[0, 1]$.

■ Now

$$f(x) = \lim_{n\to\infty} f_n(x) = \begin{cases} 0, & 0 \le x < 1 \\ 1, & x = 1 \end{cases}$$

Thus, the sequence converges pointwise to a discontinuous function on $[0, 1]$.

Let $\varepsilon > 0$ be given.

For $0 < x \le k < 1$, we have

$$|f_n(x) - f(x)| = x^n < \varepsilon$$

if $$\left(\frac{1}{x}\right)^n > \frac{1}{\varepsilon}$$

or if $$n > \log(1/\varepsilon)/\log(1/x)$$

This number, $\log(1/\varepsilon)/\log(1/x)$ increases with x, its maximum value being $\log(1/\varepsilon)/\log(1/k)$ in $]0, k]$, $k > 0$.

Let N be an integer $\ge \log(1/\varepsilon)/\log(1/k)$.

$\therefore$ $$|f_n(x) - f(x)| < \varepsilon, \quad \forall\, n \ge N, 0 < x < 1$$

Again at $x = 0$,

$$|f_n(x) - f(x)| = 0 < \varepsilon, \quad \forall\, n \ge 1$$

Thus for any $\varepsilon > 0, \exists\, N$ such that for all $x \in [0, k]$, $k < 1$

$$|f_n(x) - f(x)| < \varepsilon, \quad \forall\, n \ge N$$

Therefore, the sequence $\{f_n\}$ is uniformly convergent in $[0, k]$, $k < 1$.

However, the number $\log(1/\varepsilon)/\log(1/x) \to \infty$ as $x \to 1$ so that it is not possible to find an integer N such that $|f_n(x) - f(x)| < \varepsilon$, for all $n \ge N$ and all x in $[0, 1]$.

Hence, the sequence is not uniformly convergent on any interval containing 1 and in particular on $[0, 1]$.

Note: A point, like $x = 1$ which is such that the sequence is not uniformly convergent in any interval containing $x = 1$, is called a *point of non-uniform convergence*.

Example 3. Show that the sequence $\{f_n\}$, where

$$f_n(x) = \frac{1}{x + n}$$

is uniformly convergent in any interval $[o, b]$, $b > 0$.

- Here the sum function

$$f(x) = \lim_{n\to\infty} f_n(x) = 0 \quad \forall\ x \in [0, b]$$

so that the sequence converges pointwise to 0.

For any $\varepsilon > 0$,

$$|f_n(x) - f(x)| = \frac{1}{x+n} < \varepsilon$$

if $n > (1/\varepsilon) - x$, which decreases with x, the maximum value being $1/\varepsilon$.

Let N be an integer $\geq 1/\varepsilon$, so that for $\varepsilon > 0$, there exists N such that

$$|f_n(x) - f(x)| < \varepsilon, \quad \forall\ n \geq N$$

Hence, the sequence is uniformly convergent in any interval $[0, b]$, $b > 0$.

Example 4. Show that the sequence $\{f_n\}$, where

$$f_n(x) = \tan^{-1} nx,\ x \geq 0$$

is uniformly convergent in any interval $[a, b]$, $a > 0$, but is only pointwise convergent in $[0, b]$.

- The pointwise sum function

$$f(x) = \lim_{n\to\infty} f_n(x) = \begin{cases} \pi/2, & x > 1 \\ 0, & x = 0 \end{cases}$$

Let $\varepsilon > 0$ be given, so that for $x > 0$,

$$|f_n(x) - f(x)| = |\tan^{-1} nx - \pi/2| < \varepsilon$$

if $$\cot^{-1} nx < \varepsilon$$

or if $$n > \cot \varepsilon/x$$

which decreases when x increases, the maximum value being cot loss ε/a in $[a, b]$, $a > 0$. Let N be an integer $\geq \cot\ \varepsilon/a$.

Thus for $\varepsilon > 0, \exists\ N$ such that for all $x \in [a, b]$, $a > 0$.

$$|f_n(x) - f(x)| < \varepsilon, \quad \forall\ n \geq N$$

Hence, the sequence converges uniformly on $[a, b]$, $a > 0$.

However, $\cot\ \varepsilon/x \to \infty$ as $x \to 0$, so that no such integer N exists, such that $|f_n(x) - f(x)| < \varepsilon$, for all $n \geq N$, and hence the sequence is not uniformly convergent on $[0, b]$. It is only pointwise convergent in $[0, b]$.

Example 5. Show that the series $\sum f_n$, whose sum to n terms, $S_n(x) = nxe^{-nx^2}$ is pointwise and not uniformly convergent on any interval $[0, k]$, $k > 0$.

■ The pointwise sum $S(x) = \lim\limits_{n\to\infty} S_n(x) = 0$, for all $x \geq 0$.

Thus the series converges pointwise to 0 on $[0, k]$.

Let us suppose, if possible, the series converges uniformly on $[0, k]$, so that for any $\varepsilon > 0$, there exists an integer N such that for all $x \geq 0$,

$$|S_n(x) - S(x)| = nxe^{-nx^2} < \varepsilon, \quad \forall\, n \geq N \qquad \text{...(1)}$$

Let N_0 be an integer greater than N and $e^2\varepsilon^2$, then for $x = 1/\sqrt{N_0}$ and $n = N_0$, (1) gives

$$\sqrt{N_0}/e < \varepsilon \;\Rightarrow\; N_0 < e^2\varepsilon^2$$

so we arrive at a contradiction.

Hence, the series is not uniformly convergent on $[0, k]$.

Notes:

1. Choice of $x = 1/\sqrt{N_0}$ is admissible because the interval contains the origin.
2. The interval of uniform convergence is always be a closed interval, that is, it must include the end points. But the interval for pointwise or absolute convergence can be of any type.

Remark: These examples suggest that a discontinuity in the limit function implies a point of non-uniform convergence, although non-uniform convergence does not necessary involve discontinuity in the limit function.

Ex. 1. Show that the sequence $\{f_n\}$, where $f_n(x) = \dfrac{x}{n+x}$, is uniformly convergent in $[0, k]$, $k < \infty$ but only pointwise convergent when the interval extends to ∞.

Ex. 2. Show that the sequence $\{e^{-nx}\}$ is uniformly convergent in any interval $[a, b]$, where a and b are positive numbers but only pointwise in $[0, b]$.

Ex. 3. Show that the series $\sum f_n$, the sum of whose n terms is $S_n(x) = x/(1 + nx^2)$, converges uniformly for all real x.

Ex. 4. Show that the sequence $\left\{\dfrac{nx}{1+n^3x^2}\right\}$ converges uniformly to zero for $0 \leq x \leq 1$.

Ex. 5. Show that the sequence $\{f_n\}$, where $f_n\{x\} = \dfrac{n^2x}{1+n^3x^2}$, is not uniformly convergent on $[0, 1]$.

Ex. 6. Show that the sequences $\{nx(1-x^2)^n\}$ and $\{n^2x(1-x^2)^n\}$ are not uniformly convergent on $[0, 1]$.

Ex. 7. Show that the series

$$(1-x)^2 + x(1-x)^2 + x^2(1-x)^2 + \ldots,$$

is not uniformly convergent on $[0, 1]$.

Ex. 8. Show that the series

$$\frac{x}{1+x}+\frac{x}{(1+x)(1+2x)}+\frac{x}{(1+2x)(1+3x)}+\ldots$$

is uniformly convergent on $[a, b]$, $a > 0$ but only pointwise in $[0, b]$.

3. TESTS FOR UNIFORM CONVERGENCE

Now that we are acquainted with the meaning of the concept of uniform convergence, we shall naturally inquire how we can determine whether a given sequence or a series does or does not converge uniformly in a given interval. So far we have used merely the definition of uniform convergence for the purpose. This procedure is usually replaced by narrower tests which are more convenient in ordinary practice.

3.1 A Test for Uniform Convergence of Sequences

Theorem 3. *Let $\{f_n\}$ be a sequence of functions, such that*

$$\lim_{n\to\infty} f_n(x) = f(x),\ x \in [a, b]$$

and let

$$M_n = \operatorname*{Sup}_{x\in[a,b]} |f_n(x) - f(x)|$$

Then $f_n \to f$ uniformly on $[a, b]$ if and only if $M_n \to 0$ as $n \to \infty$.

Necessary. Let $f_n \to f$ uniformly on $[a, b]$, so that for a given $\varepsilon > 0$, there exists an integer N such that

$$|f_n(x) - f(x)| < \varepsilon,\quad \forall\, n \geq N, \forall\, x \in [a, b]$$

$$\Rightarrow \quad M_n = \operatorname*{Sup}_{x\in[a,b]} |f_n(x) - f(x)| \leq \varepsilon,\quad \forall\, n \geq N$$

$$\Rightarrow \quad M_n \to 0,\quad \text{as } n \to \infty$$

Sufficient. Let $M_n \to 0$, as $n \to \infty$, so that for any $\varepsilon > 0$, $\exists$ an integer N such that

$$M_n < \varepsilon,\ \forall\, n \geq N$$

$$\Rightarrow \quad \operatorname*{Sup}_{x\in[a,b]} |f_n(x) - f(x)| < \varepsilon,\quad \forall\, n \geq N$$

$$\Rightarrow \quad |f_n(x) - f(x)| < \varepsilon,\quad \forall\, n \geq N, \forall\, x \in [a, b]$$

$$\Rightarrow \quad f_n \to f \text{ uniformly on } [a, b].$$

Example 6. Show that the sequence $\{f_n\}$, where

$$f_n(x) = \frac{nx}{1+n^2x^2}$$

is not uniformly convergent on any interval containing zero.

■ Here

$$\lim_{n\to\infty} f_n(x) = 0, \quad \forall\ x$$

Now $\dfrac{nx}{1+n^2x^2}$ attains the maximum value $\dfrac{1}{2}$ at $x = \dfrac{1}{n}$; $\dfrac{1}{n}$ tending to 0 as $n \to \infty$. Let us take an interval $[a, b]$ containing 0.

Thus

$$M_n = \sup_{x\in[a,b]} |f_n(x) - f(x)|$$

$$= \sup_{x\in[a,b]} \left|\frac{nx}{1+n^2x^2}\right| = \frac{1}{2} \text{ which does not tend to zero as } n \to \infty.$$

Hence, the sequence $\{f_n\}$ is not uniformly convergent in any interval containing the origin.

Example 7. Prove that the sequence $\{f_n\}$, where

$$f_n(x) = \frac{x}{1+nx^2},\ x \text{ being real}$$

converges uniformly on any closed interval I.

■ Here pointwise limit,

$$f(x) = \lim_{n\to\infty} f_n(x) = 0, \quad \forall\ x$$

$$M_n = \sup_{x\in I} |f_n(x) - f(x)| = \sup_{x\in I} \left|\frac{x}{1+nx^2}\right| = \frac{1}{2\sqrt{n}} \to 0 \text{ as } n \to \infty$$

Hence $\{f_n\}$ converges uniformly on I.

$\left[\dfrac{x}{1+nx^2} \text{ attains the maximum value } \dfrac{1}{2\sqrt{n}} \text{ at } x = \dfrac{1}{\sqrt{n}}, \text{ i.e. at the origin}\right]$.

Example 8. Show that the sequence $\{f_n\}$, where

$$f_n(x) = nxe^{-nx^2},\ x \ge 0$$

is not uniformly convergent on $[0, k]$, $k > 0$

■

$$f(x) = \lim_{n\to\infty} f_n(x) = 0, \quad \forall\ x \ge 0$$

Also nxe^{-nx^2} attains maximum value $\sqrt{\dfrac{n}{2e}}$ at $x = \dfrac{1}{\sqrt{2n}}$

Now,

$$M_n = \sup_{x\in[0,k]} |f_n(x) - f(x)| = \sup_{x\in[0,k]} nxe^{-nx^2} = \sqrt{\frac{n}{2e}} \to \infty \text{ as } n \to \infty$$

Therefore, the sequence is not uniformly convergent on $[0, k]$.

Ex. 1. Show that the following sequences are not uniformly convergent on the intervals indicated:

(*i*) $\{x^n\}$ on $[0, 1]$ (*ii*) $\{e^{-nx}\}$ on $[0, k]$

Ex. 2. Test the following sequences for uniform convergence.

(*i*) $\left\{\dfrac{\sin nx}{\sqrt{n}}\right\}$, $0 \le x \le 2\pi$ (*ii*) $\left\{\dfrac{x}{n+x}\right\}$, $0 \le x \le k$

(*iii*) $\left\{\dfrac{x}{n+x}\right\}$, $0 \le x < \infty$ (*iv*) $\left\{\dfrac{n^2x}{1+n^3x^2}\right\}$, $0 \le x \le 1$

(*v*) $\left\{\dfrac{nx}{1+n^3x^2}\right\}$, $0 \le x \le 1$

3.2 Tests of Uniform Convergence of Series

Theorem 4. ***Weierstrass's M-test.*** *A series of functions* Σf_n *will converge uniformly (and absolutely) on* $[a, b]$ *if there exists a convergent series* ΣM_n *of positive numbers such that for all* $x \in [a, b]$

$$|f_n(x)| \le M_n, \quad \textit{for all } n$$

Let $\varepsilon > 0$ be a positive number.

Since ΣM_n is convergent, therefore there exists a positive integer N such that

$$|M_{n+1} + M_{n+2} + \ldots + M_{n+p}| < \varepsilon \quad \forall\, n \ge N,\ p \ge 1 \qquad \text{...(1)}$$

Hence for all $x \in [a, b]$ and for all $n \ge N$, $p \ge 1$, we have

$$|f_{n+1}(x) + f_{n+2}(x) + \ldots + f_{n+p}(x)| \le |f_{n+1}(x)| + |f_{n+2}(x)| + \ldots + |f_{n+p}(x)| \qquad \text{...(2)}$$

$$\le M_{n+1} + M_{n+2} + \ldots + M_{n+p} < \varepsilon \qquad \text{...(3)}$$

Equations (2) and (3) imply that Σf_n is uniformly and absolutely convergent on $[a, b]$.

Remarks:

1. The converse is not asserted and is in fact, not true *i.e.*, non-convergence of ΣM_n does not imply anything as far as Σf_n is concerned.
2. Series which satisfy the *M*-test have been called *normally convergent* by *Baire*, to emphasize the fact that such series are uniformly as well as absolutely convergent. The terminology has the additional advantage of emphasising the fact that the test can be applied to nearly all series in ordinary everyday use.

ILLUSTRATIONS

1. The series $\Sigma r^n \cos n\theta$, $\Sigma r^n \sin n\theta$, $\Sigma r^n \cos n^2\theta$, $\Sigma r^n \sin(a^n\theta)$, $0 < r < 1$, converge uniformly for all real values of θ.

   The result follows by taking $M_n = r^n$.

2. The series $\Sigma \dfrac{a_n x^n}{1+x^{2n}}$, $\Sigma \dfrac{a_n x^{2n}}{1+x^{2n}}$ converge uniformly for all real values of x, if Σa_n is absolutely convergent.

3. $\Sigma \dfrac{\sin(x^2 + n^2 x)}{n(n+1)}$ is uniformly convergent for all real x.

$\left[\text{Take } M_n = \dfrac{1}{n(n+1)}\right]$.

4. $\Sigma \dfrac{\cos n\theta}{n^p}$ is uniformly and absolutely convergent for all real values of $\theta, p > 1$.

$\left[\text{Take } M_n = \dfrac{1}{n^p}\right]$.

5. $\Sigma \dfrac{(-1)^n x^{2n}}{n^p(1 + x^{2n})}$ converges absolutely and uniformly for all real x if $p > 1$.

$\left[\text{Take } M_n = \dfrac{1}{n^p}\right]$.

6. Σn^{-x} is uniformly convergent in $[1 + \delta, \infty[, \delta > 0$.

Example 9. Test for uniform convergence, the series

$$\frac{2x}{1+x^2} + \frac{4x^3}{1+x^4} + \frac{8x^7}{1+x^8} + \ldots, -\frac{1}{2} \le x \le \frac{1}{2}$$

■ The nth term $f_n(x) = \dfrac{2^n x^{2n-1}}{1 + x^{2n}}$

$$| f_n(x) | \le 2^n (\alpha)^{2n-1}$$

where $|x| \le \alpha \le \dfrac{1}{2}$.

The series $\Sigma 2^n (a)^{2n-1}$ converges, and hence by M-test the given series converges uniformly on $\left[-\dfrac{1}{2}, \dfrac{1}{2}\right]$.

Example 10. Show that the series $\Sigma \dfrac{x}{n^p + x^2 n^q}$ converges uniformly over any finite interval $[a, b]$, for

(*i*) $p > 1, q \ge 0$ (*ii*) $0 < p \le 1, p + q > 2$

■ (*i*) When $p > 1, q \ge 0$

$$|f_n(x)| = \left|\frac{x}{n^p + x^2 n^q}\right| \le \frac{\alpha}{n^p}$$

where $\alpha \ge \max\{|a|, |b|\}$.

The series $\Sigma(\alpha/n^p)$ converges for $p > 1$.

Hence by *M*-test, the given series converges uniformly over the interval $[a, b]$.

(*ii*) When $0 < p \le 1,\ p + q > 2$.

$|f_n(x)|$ attains the maximum value $\dfrac{1}{2n^{\frac{1}{2}(p+q)}}$ at the point, where $x^2 n^q = n^p$

$$\therefore \qquad |f_n(x)| \le \frac{1}{2n^{\frac{1}{2}(p+q)}}$$

The series $\Sigma \dfrac{1}{2n^{\frac{1}{2}(p+q)}}$ converges for $p+q>2$. Hence by *M*-test, the given series converges uniformly over any finite interval $[a, b]$.

In spite of its great practical importance, Weierstrass's *M*-test is necessarily applicable to a restricted class of series—series which are absolutely convergent as well. When this is not the case, we have to make use of more delicate tests, which we construct by analogy with those for series of arbitrary terms—*Abel's* and *Dirichlet's* test.

Theorem 5. ***Abel's test.*** *If $b_n(x)$ is a positive, monotonic decreasing function of n for each fixed value of x in the interval $[a, b]$, and $b_n(x)$ is bounded for all values of n and x concerned, and if the series $\Sigma u_n(x)$ is uniformly convergent on $[a, b]$, then so also is the series $\Sigma b_n(x)u_n(x)$.*

Since $b_n(x)$ is bounded for all values of n and for x in $[a, b]$, therefore there exists a number $K > 0$, independent of x and n, such that for all $x \in [a, b]$,

$$0 \le b_n(x) \le K, \quad (\text{for } n = 1, 2, 3, \ldots) \qquad \ldots(1)$$

Again, since $\Sigma u_n(x)$ converges uniformly on $[a, b]$, therefore for any $\varepsilon > 0$, we can find an integer N such that

$$\left| \sum_{r=n+1}^{n+p} u_r(x) \right| < \frac{\varepsilon}{K}, \quad \forall\, n \ge N,\ p \ge 1 \qquad \ldots(2)$$

Hence using Abel's lemma (Ch. 9, § 13) we get

$$\left| \sum_{r=n+1}^{n+p} b_r(x) u_r(x) \right| \le b_{n+1}(x) \max_{q=1,2,\ldots,p} \left| \sum_{r=n+1}^{n+q} u_r(x) \right|$$

$$< K \frac{\varepsilon}{K} = \varepsilon, \text{ for } n \ge N,\ p \ge 1,\ a \le x \le b$$

$\Rightarrow$ $\Sigma b_n(x)u_n(x)$ is uniformly convergent on $[a, b]$.

Corollary 1. *A uniformly convergent series $\Sigma u_n(x)$ remains uniformly convergent on $[a, b]$, if its each term is multiplied by a function $a_n(x)$, $a \le x \le b$, provided that the sequence $\{a_n(x)\}$ is uniformly bounded on $[a, b]$ (i.e., $\exists\, K > 0$, such that $|a_n(x)| \le K$, for all x in $[a, b]$ and for all n), and monotonic in n, for each $x \in [a, b]$.*

Under the given conditions, $\{a_n(x)\}$ converges pointwise. Let us write for each $x \in [a, b]$,

$$b_n(x) = \{\lim_{n\to\infty} a_n(x)\} - a_n(x), \text{ or } a_n(x) - \lim_{n\to\infty} a_n(x),$$

according as $\{a_n(x)\}$ is monotonic increasing or decreasing. With this function $b_n(x)$, we deduce as above that the series $\Sigma\, b_n(x)\, u_n(x)$ converges uniformly on $[a, b]$. Also, since $\Sigma\, u_n(x)$ and hence $\Sigma\,[\lim a_n(x)]\;\; u_n(x)\;\; (\because |\lim a_n(x)| \le K, a \le x \le b)$ converges uniformly on $[a, b]$. The uniform convergence of $\Sigma\, a_n(x)\, u_n(x)$, then follows easily.

Corollary 2. *If* $\sum_{n=1}^{\infty} a_n x^n$ *is a (power) series which converges for all values of x, where* $|x| < R$, *then* $\Sigma a_n x^n$ *is uniformly convergent in* $[0, R]$ *if and only if* $\Sigma a_n R^n$ *is convergent.*

Let $\Sigma a_n R^n$, be convergent, so that being a series of real numbers, it is uniformly convergent in $[0, R]$.

Now, since $\Sigma a_n R^n$ is uniformly convergent, and $(x/R)^n$ is a positive monotonic decreasing bounded function of n, for each value of x in $[0, R]$, therefore by Abel's test, the series $\Sigma a_n\, R^n (x/R)^n < \Sigma a_n x^n$ is uniformly convergent on $[a, b]$.

If the series $\Sigma a_n x^n$ is uniformly convergent in $[0, R]$, it is obviously convergent at $x = R$.

Example 11. The series $\Sigma \dfrac{(-1)^n}{n} |x|^n$ is uniformly convergent in $-1 \le x \le 1$.

- Since $|x|^n$ is positive, monotonic decreasing and bounded for $-1 \le x \le 1$, and the series $\Sigma \dfrac{(-1)^n}{n}$ is uniformly convergent, therefore $\Sigma \dfrac{(-1)^n}{n} |x|^n$ is also so in $-1 \le x \le 1$.

Example 12. Show that $\Sigma\, a_n / n^x$ converges uniformly in $[0, 1]$, if $\Sigma\, a_n$ converges.

- Since $1/n^x$ is a positive, monotonic decreasing function and is bounded by 0 and 1 in $[0, 1]$, and the series $\Sigma\, a_n$ is convergent (and so uniformly), therefore $\Sigma\, a_n / n^x$ is uniformly convergent in $[0, 1]$.

Ex. If $\Sigma\, a_n$ is convergent, then show that each of the following series is uniformly convergent in $[0, 1]$.

$$\Sigma a_n x^n,\; \Sigma a_n \frac{x^n}{1 + x^n},\; \Sigma \frac{a_n x^n}{1 + x^{2n}},\; \Sigma \frac{n x^n (1 - x)}{1 + x^n} a_n,\; \Sigma \frac{2n a_n x^n (1 - x)}{1 + x^{2n}}$$

Theorem 6. ***Dirichlet's test.*** *If* $b_n(x)$, *is a monotonic function of n for each fixed value of x in* $[a, b]$, *and* $b_n(x)$ *tends uniformly to zero for* $a \le x \le b$, *and if there is a number* $K > 0$ *independent of x and n, such that for all values of x in* $[a, b]$,

$$\left| \sum_{r=1}^{n} u_r(x) \right| \le K,\;\; \forall\, n$$

then the series $\Sigma\, b_n(x) u_n(x)$ *is uniformly convergent on* $[a, b]$.

First method. We may assume that $b_n(x)$ is a positive monotonic decreasing function of n, for each $x \in [a, b]$, since the general case follows by the procedure given in the above cor. 1. Now $b_n(x)$ tends uniformly to zero, therefore for any $\varepsilon > 0$, we can find an integer N (independent of x) such that for all values of x in $[a, b]$.

$$0 \le b_n(x) < \varepsilon/2K, \text{ for all } n \ge N$$

For such values of n and any integral value of $p \ge 1$, we have by Abel's Lemma,

$$\left|\sum_{r=n+1}^{n+p} b_r(x)u_r(x)\right| \le b_{n+1}(x) \max_{q=1,2,\ldots,p} \left|\sum_{r=n+1}^{n+q} u_r(x)\right|$$

$$\le b_{n+1}(x)\left\{\left|\sum_{r=1}^{n} u_r(x)\right| + \max_{q=1,2,\ldots,p}\left|\sum_{r=n}^{n+q} u_r(x)\right|\right\}$$

$$< \frac{\varepsilon}{2K}(K+K) = \varepsilon$$

Hence by Cauchy's criterion the series $\sum b_n(x)u_n(x)$ converges uniformly for $x \in [a, b]$.

Second method. Since $b_n(x)$ tends uniformly to zero, therefore for any $\varepsilon > 0$, there exists an integer N (independent of x) such that for all $x \in [a, b]$,

$$|b_n(x)| < \varepsilon/4K, \text{ for all } n \ge N.$$

Let $S_n(x) = \sum_{r=1}^{n} u_r(x)$, for all $x \in [a, b]$, and for all n,

$$\therefore \quad \sum_{r=n+1}^{n+p} b_r(x)u_r(x) = b_{n+1}(x)\{S_{n+1} - S_n\} + b_{n+2}(x)\{S_{n+2} - S_{n+1}\} + \ldots$$

$$+ b_{n+p}(x)\{S_{n+p} - S_{n+p-1}\}$$

$$= -b_{n+1}(x)S_n + \{b_{n+1}(x) - b_{n+2}(x)\}S_{n+1} + \ldots$$

$$+ \{b_{n+p-1}(x) - b_{n+p}(x)\}S_{n+p-1} + b_{n+p}(x)S_{n+p}$$

$$= \sum_{r=n+1}^{n+p-1}\{b_r(x) - b_{r+1}(x)\}S_r(x) - b_{n+1}(x)S_n(x)$$

$$+ b_{n+p}(x)S_{n+p}(x)$$

$$\left|\sum_{r=n+1}^{n+p} b_r(x)u_r(x)\right| \le \sum_{r=n+1}^{n+p-1} |b_r(x) - b_{r+1}(x)||S_r(x)| + |b_{n+1}(x)||S_n(x)| +$$

$$+ |b_{n+p}(x)||S_{n+p}(x)|$$

Making use of the monotonicity of $b_n(x)$

$$\sum_{r=n+1}^{n+p-1} |b_r(x) - b_{r+1}(x)| = |b_{n+1}(x) - b_{n+p}(x)|, \text{ for } a \le x \le b,$$

and the relation $|S_n(x)| \le K$, for all $x \in [a, b]$ and for all $n = 1,2,3,\ldots$, we deduce that for all $x \in [a, b]$ and all $p \ge 1, n \ge N$

$$\left|\sum_{r=n+1}^{n+p} b_r(x)u_r(x)\right| \le K\,|b_{n+1}(x) - b_{n+p}(x)| + \frac{\varepsilon}{4K}\,2K$$

$$< K\,\frac{\varepsilon}{2K} + \frac{\varepsilon}{2} = \varepsilon.$$

Hence by Cauchy's criterion, the series $\sum b_n(x)u_n(x)$ converges uniformly on $[a, b]$.

Remark: The statement '$\left|\sum_{r=1}^{n} u_r(x)\right| \le k \quad \forall\, x \in [a, b]$, and for all n' amounts to saying that the sequences of partial sums of $\sum u_n(x)$ are bounded for each value of $x \in [a, b]$, *i.e.*, for each point $x_i \in [a, b]$ there is a number K_i such that $\left|\sum_{r=1}^{n} u_r(x)\right| \le K_i$, and there exists a number K such that $K_t < K, \ \forall\, i$.

This fact is expressed by saying that 'the partial sums of the series are *uniformly bounded*'.

This in turn amounts to saying that 'the series $\sum u_n(x)$ either converges uniformly or oscillates finitely'.

So, *Dirichlet's test* can be stated also as:

If $b_n(x)$ is a monotonic function of n for each fixed value of x in $[a, b]$, and $b_n(x)$ tends uniformly to zero for $a \le x \le b$, and if $\sum u_n(x)$ either uniformly converges or oscillates finitely in $[a, b]$, then the series $\sum b_n(x)u_n(x)$ is uniformly convergent on $[a, b]$.

Example 13. Prove that the series $\sum(-1)^n \dfrac{x^2 + n}{n^2}$, converges uniformly in every bounded interval, but does not converge absolutely for any value of x.

- Let the bounded interval be $[a, b]$, so that $\exists$ a number K such that for all x in $[a, b]$, $|x| < K$.

  Let us take $\sum u_n = \sum(-1)^n$ which oscillates finitely and

$$b_n = \frac{x^2 + n}{n^2} < \frac{K^2 + n}{n^2}$$

Clearly b_n is a positive, monotonic decreasing function of n for each x in $[a, b]$, and tends to zero uniformly for $a \le x \le b$.

Hence by Dirichlet's test, the series $\sum(-1)^n \dfrac{x^2 + n}{n^2}$ converges uniformly on $[a, b]$.

Again $\sum \left|(-1)^n \frac{x^2+n}{n^2}\right| = \sum \frac{x^2+n}{n^2} \sim \sum \frac{1}{n}$, which diverges. Hence the given series is not absolutely convergent for any value of x.

Example 14. Prove that the series $\sum \frac{\cos n\theta}{n^p}$ and $\sum \frac{\sin n\theta}{n^p}$ converge uniformly for all values of $p > 0$ in an interval $[\alpha, 2\pi - \alpha[$, where $0 < \alpha < \pi$.

■ When $p > 1$, Weierstrass's *M*-test at once proves that both the series converge uniformly for all real values of θ.

When $0 < p \le 1$, both the series converge uniformly in any interval $[\alpha, 2\pi - \alpha]$, $\alpha > 0$. This can be proved by taking $b_n = (1/n^p)$ and $u_n = \cos n\theta$ (or $\sin n\theta$) in Dirichlet's test.

$(1/n^p)$ is positive monotonic decreasing and tends uniformly to zero for $0 < p \le 1$, and

$$\left|\sum_{r=1}^{n} u_r\right| = \left|\sum_{r=1}^{n} \cos r\theta\right| = |\cos\theta + \cos 2\theta + \ldots + \cos n\theta|$$

$$= \left|\frac{\cos((n+1)/2)\theta \sin(n/2)\theta}{\sin(\theta/2)}\right| \le \operatorname{cosec}(\alpha/2), \ \forall\, n,$$

for $\theta \in [\alpha, 2\pi - \alpha]$,

Thus, all the conditions are fulfilled and the series $\sum(\cos n\theta/n^p)$ and similarly $\sum(\sin n\theta/n^p)$ converge uniformly on $[\alpha, 2\pi - \alpha]$ where $0 < \alpha < \pi$.

Example 15. Show that the series $\sum\{\log(n+1)\}^{-x} \cos nx$ is uniformly convergent in $[\theta_1, \theta_2]$, where $0 < \theta_1 \le x \le \theta_2 < 2\pi$.

■ When $x \in [\theta_1, \theta_2]$, $\{\log(n+1)\}^{-x}$ is a positive monotonic decreasing function of n. Also since $\{\log(n+1)\}^{-x} \le \{\log(n+1)\}^{-\theta_1}$, the function $\{\log(n+1)\}^{-x}$ tends uniformly to zero as $n \to \infty$. Moreover, as in Example 12.14,

$$\left|\sum_{r=1}^{n} \cos rx\right| \le \frac{1}{\sin(x/2)} \le \max\left(\frac{1}{\sin(\theta_1/2)}, \frac{1}{\sin(\theta_2/2)}\right),$$

both are independent of x and n.

Thus by Dirichlet's test the series $\sum\{\log(n+1)\}^{-x} \cos nx$ is uniformly convergent in $[\theta_1, \theta_2]$.

Notes:

1. It is to be understood that θ_1 can be as close to zero and θ_2 as close to 2π as we please.
2. If $\{v_n\}$ is a monotonic sequence of real numbers that converges to zero, then each of the series $\sum v_n \sin n\theta$, $\sum v_n \cos n\theta$ is uniformly convergent with regard to θ in the interval $[\alpha, 2\pi - \alpha]$ where α is any fixed positive number less than π.

EXERCISE

1. Prove that the series,

$$(i)\ \Sigma \frac{x^n}{n^2}, \qquad (ii)\ \Sigma \frac{x^n}{n(n+1)}, \qquad (iii)\ \Sigma \frac{x^{2n}}{n^2 + x^{2n}}$$

are uniformly convergent in $[-1, 1]$.

2. Prove that, if α is any fixed positive number less than unity, each of the series

$$\Sigma x^n,\ \Sigma (n+1)^{-1} x^n,\ \Sigma (n+1)\, x^n,\ \Sigma n^3 x^n$$

is uniformly convergent in $[-\alpha, \alpha]$.

3. (*a*) Show that each of the series

$$\Sigma \frac{1}{n^4 + n^2 x^2},\ \Sigma \frac{1}{n^2 + n^4 x^2}$$

is uniformly convergent in $[-k, k]$, for real k.

(*b*) Show that $\Sigma\, 1/n^x$ converges uniformly for all real $x > 1$.

(*c*) Show that the series

$$1 + \frac{e^{-2x}}{2^2 - 1} - \frac{e^{-4x}}{4^2 - 1} + \frac{e^{-6x}}{6^2 - 1} - \dots$$

converges uniformly for all real $x \geq 0$.

4. Show that the series $\sum\limits_{1}^{\infty} \dfrac{x}{n(1 + nx^2)}$ converges uniformly for all real x.

[*Hint*: The maximum value of $f_n(x)$ is $1/2n^{3/2}$ at $x^2 = 1/n$, apply *M*-test.]

5. Prove that each of the series, $\Sigma \dfrac{\sin nx}{n}, \Sigma \dfrac{\cos nx}{n}$ converges uniformly with respect to x *in* $[\alpha, 2\pi - \alpha]$, where α is any fixed positive number less than π.

Prove also that each of the series, $\Sigma \dfrac{\sin nx}{n^2}, \Sigma \dfrac{\cos nx}{n^2}$ is uniformly convergent in $[0, 2\pi]$.

6. Show that the series

$$\frac{1}{a} - \frac{2a}{a^2 - 1^2} \cos\theta + \frac{2a}{a^2 - 2^2} \cos 2\theta - \dots$$

is uniformly convergent with respect to θ in any finite interval.

$\left[\textit{Hint}: |a^2 - n^2| = |n^2 - a^2| > n^2/2, \text{ when } n \text{ exceeds a certain number } N, \text{ so that } \left|\dfrac{\cos n\theta}{a^2 - n^2}\right| \leq \dfrac{2}{n^2}\right].$

7. Discuss the series $\Sigma(-x)^n / n(1 + x^n)$ for uniform convergence for real x.

8. Show that $\Sigma \dfrac{\log n}{n^x}$ converges uniformly for all real $x \geq 1 + \alpha > 1$.

$\left[\textit{Hint}: |(\log n)/n^x| \leq \dfrac{1}{n^{1+\alpha/2}} \cdot \dfrac{\log n}{n^{\alpha/2}} < \dfrac{1}{n^{1+\alpha/2}}.\right]$

9. Discuss the uniform convergence with respect to x of the series, $\sum\limits_{n=1}^{\infty} \dfrac{(-1)^n}{\sqrt{n}} \sin\left(1 + \dfrac{x}{n}\right)$, over any closed and bounded subset of $\mathbf{R}$.

4. PROPERTIES OF UNIFORMLY CONVERGENT SEQUENCES AND SERIES

Whereas we saw earlier that fundamental properties of the functions f_n do not in general hold for the pointwise limit function f, we shall now show that roughly speaking these properties hold for the limit function f when the convergence is uniform. In this connection we now give some theorems which become particularly important in application.

4.1. Theorem 7(A). *If a sequence $\{f_n\}$ converges uniformly in $[a, b]$ and x_0 is a point of $[a, b]$ such that*

$$\lim_{x \to x_0} f_n(x) = a_n, \; n = 1, 2, 3, \ldots$$

then (*i*) *$\{a_n\}$ converges,*

and (*ii*) $\lim_{x \to x_0} f(x) = \lim_{n \to \infty} a_n$

[The existence of $\lim_{x \to x_0} f(x)$ is a part of conclusion].

(*i*) Since the sequence $\{f_n\}$ converges uniformly on $[a, b]$, therefore for $\varepsilon > 0$, there exists an integer m (independent of x) such that for all $x \in [a, b]$,

$$| f_{n+p}(x) - f_n(x) | < \varepsilon/2, \quad \forall \, n \geq m, \, p \geq 1$$

Keeping n, p fixed and letting $x \to x_0$, we get

$$| a_{n+p} - a_n | \leq \varepsilon/2 < \varepsilon, \quad \forall \, n \geq m, \, p \geq 1$$

$\Rightarrow$ the sequence $\{a_n\}$ converges, say to A. ...(1)

(*ii*) Since $\{f_n\}$ converges uniformly to f, therefore for any $\varepsilon > 0$, there exists an integer N_1 such that for all $x \in [a, b]$,

$$| f_n(x) - f(x) | < \varepsilon/3, \quad \forall \, n \geq N_1 \qquad \text{...(2)}$$

Similarly there exists an integer N_2, such that

$$| a_n - A | < \varepsilon/3, \quad \forall \, n \geq N_2 \qquad \text{...(3)}$$

Let $N = \max(N_1, N_2)$.

Again, since $\lim_{x \to x_0} f_n(x) = a_n$, for all n, therefore $\lim_{x \to x_0} f_N(x) = a_N$ and so for $\varepsilon > 0$, there exists a $\delta > 0$, such that for $| x - x_0 | < \delta$, we have

$$| f_N(x) - a_N | < \varepsilon/3, \qquad \text{...(4)}$$

Hence for $| x - x_0 | < \delta$, we have

$$| f(x) - A | \leq | f(x) - f_N(x) | + | f_N(x) - a_N | + | a_N - A |$$

$$< \varepsilon/3 + \varepsilon/3 + \varepsilon/3 = \varepsilon \quad \text{[using (2), (3), (4)]}$$

$\Rightarrow$ $\lim_{x \to x_0} f(x)$ exists and equals A

Thus, $\lim_{x \to x_0} f(x) = \lim_{n \to \infty} a_n$

or equivalently

$$\lim_{x \to x_0} \lim_{n \to \infty} f_n(x) = \lim_{n \to \infty} \lim_{x \to x_0} f_n(x)$$

Theorem 7(B). *If a series $\sum_{n=1}^{\infty} f_n$ converges uniformly to f in $[a, b]$, and x_0 is a point in $[a, b]$ such that*

$$\lim_{x \to x_0} f_n(x) = a_n, \ (n = 1, 2, 3, \ldots)$$

then (*i*) $\sum_{n=1}^{\infty} a_n$ *converges,*

and (*ii*) $\lim_{x \to x_0} f(x) = \sum_{n=1}^{\infty} a_n$

[The existence of $\lim_{x \to x_0} f(x)$ is a part of conclusion.]

(*i*) Since the series $\sum f_n$ converges uniformly on $[a, b]$, for $\varepsilon > 0$, $\exists$ an integer m such that for all $x \in [a, b]$ and for any integer p,

$$\left| \sum_{r=n+1}^{n+p} f_r(x) \right| < \varepsilon/2, \ \ \forall \ n \geq m, \ p \geq 1$$

Keeping n, p fixed and letting $x \to x_0$, we get

$$\left| \sum_{r=n+1}^{n+p} a_r \right| \leq \varepsilon/2 < \varepsilon, \ \ \forall \ n \geq m, \ p \geq 1$$

$\Rightarrow$ The series $\sum a_n$ converges to A. ...(1)

(*ii*) Since $\sum f_n$ converges uniformly to f, therefore for $\varepsilon > 0$, $\exists \ N_1$ such that for all $x \in [a, b]$,

$$\left| \sum_{r=1}^{n} f_r(x) - f(x) \right| < \frac{\varepsilon}{3}, \ \ \forall \ n \geq N_1 \quad \text{...(2)}$$

Similarly,

$$\left| \sum_{r=1}^{n} a_r - A \right| < \frac{\varepsilon}{3}, \ \ \forall \ n \geq N_2 \quad \text{...(3)}$$

Let $N = \max\ (N_1, N_2)$

Again, since

$$\lim_{x \to x_0} f_n(x) = a_n, \ n = 1, 2, 3, \ldots$$

Therefore for the above $\varepsilon > 0$, it is possible to choose $\delta > 0$ such that for $n = 1, 2, 3, \ldots N$, we have (taking $\delta = \min\ \{\delta_1, \delta_2, \ldots, \delta_N\}$.)

$$|f_n(x) - a_n| < \frac{\varepsilon}{3N}, \text{ for } |x - x_0| < \delta$$

$$\therefore \qquad \left| \sum_{r=1}^{N} f_r(x) - \sum_{r=1}^{N} a_r \right| \le \sum_{r=1}^{N} | f_r(x) - a_r |$$

$$< N \cdot \frac{\varepsilon}{3N} = \frac{\varepsilon}{3}, \text{ for } | x - x_0 | < \delta \qquad \ldots(4)$$

Hence for $| x - x_0 | < \delta$, we have

$$| f(x) - A | \le \left| f(x) - \sum_{r=1}^{N} f_r(x) \right| + \left| \sum_{r=1}^{N} f_r(x) - \sum_{r=1}^{N} a_r \right| + \left| \sum_{r=1}^{N} a_r - A \right|$$

$$< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon, \text{ (Using 2, 3, 4)}$$

$\Rightarrow \qquad \lim_{x \to x_0} f(x)$ exists and equals A.

CHAPTER 12

Remark: The result simply states, "the limit of the sum function of a series = the sum of the series of limits of functions", *i.e.*,

$$\lim_{x \to x_0} \sum_{n=1}^{\infty} f_n(x) = \sum_{n=1}^{\infty} \lim_{x \to x_0} f_n(x)$$

We now prove a theorem which though of great significance appears to be a special case of the theorem proved above.

4.2 Uniform Convergence and Continuity

Theorem 8(A). *If $\{f_n\}$ is a sequence of continuous functions on an interval $[a, b]$, and if $f_n \to f$ uniformly on $[a, b]$, then f is continuous on $[a, b]$.*

(B). *If a series $\sum f_n$ converges uniformly to f in an interval $[a, b]$ and its terms f_n are continuous at a point x_0 of the interval, then the sum function f is also continuous at x_0.*

(B). Since $\sum f_n$ converges uniformly to f on $[a, b]$, therefore for $\varepsilon > 0$, we can choose N such that for all x in $[a, b]$.

$$\left| \sum_{r=1}^{n} f_r(x) - f(x) \right| < \frac{\varepsilon}{3}, \quad \forall\, n \ge N \qquad \ldots(1)$$

and in particular, at a point x_0 in $[a, b]$, and $n = N$

$$\left| \sum_{r=1}^{N} f_r(x_0) - f(x_0) \right| < \frac{\varepsilon}{3}, \qquad \ldots(2)$$

Again, since each f_n is continuous at x_0, the sum of a finite number of functions, $\sum_{r=1}^{N} f_r$ is also continuous at $x = x_0$.

Therefore for $\varepsilon > 0$, $\exists\, \delta > 0$, such that

$$\left| \sum_{r=1}^{N} f_r(x) - \sum_{r=1}^{N} f_r(x_0) \right| < \frac{\varepsilon}{3}, \quad \text{for } |x - x_0| < \delta \qquad \ldots(3)$$

Hence for $|x - x_0| < \delta$, we have

$$|f(x) - f(x_0)| \le \left| f(x) - \sum_{r=1}^{N} f_r(x) \right| + \left| \sum_{r=1}^{N} f_r(x) - \sum_{r=1}^{N} f_r(x_0) \right|$$

$$+ \left| \sum_{r=1}^{N} f_r(x_0) - f(x_0) \right|$$

$$< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon \quad \text{[using (1), (2) \& (3)]}$$

$$\Rightarrow \qquad f(x) \to f(x_0) \quad \text{when } x \to x_0$$

i.e., the sum function f is continuous at $x = x_0$.

Corollary. Since x_0 is an arbitrary point of $[a, b]$, the theorem holds for all points of $[a, b]$, and so we state:

"If a series $\sum f_n$ converges uniformly to f on an interval and if the functions f_n are all continuous throughout the interval, then so is the sum function f."

Remarks:

1. The converse of the theorem is neither asserted nor it is true, *i.e.*, series (sequence) of continuous terms exist which have a continuous sum (limit) but are not uniformly convergent. In other words, *the condition of uniform convergence is only sufficient, not necessary.*
2. *However, if the sum function (limit function) of a series (sequence) of continuous terms is not continuous on an interval, the convergence cannot be uniform.* This conclusion is very often used with the advantage in deciding that the convergence is not uniform.

ILLUSTRATIONS

1. The sequence $\{x^n\}$ or $\{\tan^{-1} nx\}$ of continuous functions has a discontinuous limit function on $[0, 1]$. Therefore, the convergence is not uniform on $[0, 1]$.
2. The sequence $\{nx/(1 + n^2x^2)\}$ or $\{nxe^{-nx^2}\}$ of continuous function has a continuous limit function on $[0, 1]$, although the convergence is not uniform.
3. The sum function $(1 + x)$ of the series $\sum(1 - x^2)x^n$ is continuous on $[0, 1]$ although the convergence is not uniform.
4. The sequence $\{1/(x + n)\}$ converges uniformly to the continuous function 0 for all real $x \ge 0$.
5. The sum function of the series $\sum_{n=0}^{\infty} (1 - x)x^n$ is $f(x) = \begin{cases} 1, & x \ne 1 \\ 0, & x = 1 \end{cases}$ which is discontinuous on $[0, 1]$. Therefore, the series is not uniformly convergent on $[0, 1]$.

Note: There is a special class of sequences (series) for which uniform convergence is equivalent to the continuity of the limit (sum) function of the sequence (series). In that connection, we do a theorem, due to **Dini**, an Italian mathematician.

Theorem 9. ***Dini's Theorem on uniform convergence.*** **(A).** *If a sequence of continuous functions $\{f_n\}$, defined on $[a, b]$ is monotonic increasing, and converges (pointwise) to a continuous function f, then the convergence is uniform on $[a, b]$.*

(B). *If the sum function of a series $\sum f_n$, with non-negative continuous terms defined on an interval $[a, b]$ is continuous on $[a, b]$, then the series is uniformly convergent on the interval.*

(A). Since the sequence $\{f_n\}$ is monotonic increasing, and converges to f on $[a, b]$, therefore, for any $\varepsilon > 0$ and each x in $[a, b]$ there is an integer N such that

$$0 \leq f(x) - f_n(x) < \varepsilon \qquad \text{...(1)}$$

Let $R_n(x) = f(x) - f_n(x), \ n = 1, 2, 3, \ldots$

Clearly the sequence $\{R_n\}$ is monotonic decreasing, *i.e.*

$$R_1(x) \geq R_2(x) \geq \ldots \geq R_n(x) \geq \ldots \qquad \text{...(2)}$$

and is bounded below by zero.

Thus, the sequence $\{R_n\}$ converges pointwise to zero on $[a, b]$.

However, if (1) and (2) hold for all x in $[a, b]$ and N is independent of x, then the convergence is uniform.

Suppose, if possible, that for a certain $\varepsilon_0 > 0$, no such N independent of x exists. Then for each $n = 1, 2, \ldots$, there is $x_n \in [a, b]$ such that

$$R_n(x_n) \geq \varepsilon_0 \qquad \text{...(3)}$$

The sequence $\{x_n\}$ of points belonging to the interval $[a, b]$ is bounded and therefore has at least *one limit point* ξ in $[a, b]$.

Consequently we can assert that there is a subsequence $\{x_{n_k}\}$ of $\{x_n\}$, convergent to ξ *i.e.*, $x_{n_k} \to \xi$ as $k \to \infty$.

The function $R_n(x) = f(x) - f_n(x)$ being the difference of two continuous functions is continuous and therefore, we can write, for every fixed m, the relation

$$\lim_{k \to \infty} R_m(x_{n_k}) = R_m(\xi)$$

But for every m and any sufficiently large k we have $n_k \geq k > m$ and consequently, in view of (2) and (3), we get

$$R_m(x_{n_k}) \geq R_{n_k}(x_{n_k}) \geq \varepsilon_0$$

CHAPTER 12

Proceeding to limits as $k \to +\infty$, we see that $R_m(\xi) \geq \varepsilon_0$, for any m, which contradicts the relation $\lim_{m\to\infty} R_m(\xi) = 0$, implied by the pointwise convergence of the sequence $\{R_n\}$ on $[a, b]$. Hence the theorem.

(B). The partial sums $S_n(x) = \sum_{r=1}^{n} f_r(x)$, with non-negative continuous terms f_r, form a non-decreasing sequence of continuous functions convergent point wise to a continuous function f. Therefore by (A), the sequence converges uniformly and thus the series is also uniformly convergent.

Example 16. Show that the series

$$x^4 + \frac{x^4}{1+x^4} + \frac{x^4}{(1+x^4)^2} + \frac{x^4}{(1+x^4)^3} + \ldots$$

is not uniformly convergent on [0, 1].

■ The terms of the series are continuous and the series converges pointwise of f, where

$$f(x) = \begin{cases} 1 + x^4, & x \neq 0 \\ 0, & x = 0 \end{cases}$$

which is a discontinuous function on [0, 1].

Hence the series cannot converge uniformly on [0, 1].

Example 17. Show that the series $\sum \frac{x}{(nx+1)\,\{(n-1)x+1\}}$, is uniformly convergent on any interval, $[a, b]$, $0 < a < b$, but only pointwise on $[0, b]$.

■ Let

$$f_n(x) = \frac{x}{(nx+1)\,\{(n-1)x+1\}}$$

$$= \frac{1}{(n-1)x+1} - \frac{1}{nx+1}$$

$\therefore$ nth partial sum $S_n(x) = \sum_{r=1}^{n} f_r(x) = 1 - \frac{1}{nx+1}$

$\therefore$ The sum function $f(x) = \begin{cases} 1, & x \neq 0 \\ 0, & x = 0 \end{cases}$

Thus f is discontinuous on $[0, b]$ and therefore convergence is not uniform on $[0, b]$, it is only pointwise.

When $x \neq 0$, let $\varepsilon > 0$ be given

$$|S_n(x) - f(x)| = \frac{1}{nx+1} < \varepsilon$$

When $n > \frac{1}{x}\left(\frac{1}{\varepsilon} - 1\right)$, but $\frac{1}{x}\left(\frac{1}{\varepsilon} - 1\right)$ decreases with x, let its maximum value $\frac{1}{a}\left(\frac{1}{\varepsilon} - 1\right) = m_0$ (independent of x) on $[a, b]$.

Thus for all $x \in [a, b]$, $\exists$ an integer m $(> m_0)$, such that

$$|S_n(x) - f(x)| < \varepsilon, \quad \text{for } n \geq m$$

Hence, the series converges uniformly on $[a, b]$, $0 < a < b$.

Example 18. Show that the series $\sum_{1}^{\infty} \frac{(-1)^{n-1}}{n + x^2}$ is uniformly convergent but not absolutely for all real values of x.

■ The given series converges by Leibnitz Test. However $\Sigma \frac{1}{n + x^2}$ behaves as $\Sigma \frac{1}{n}$ and is therefore divergent. Hence the series is not absolutely convergent for any value of x.

Again, let $S_n(x)$ denotes the nth partial sum and S(x), the sum of the series

$$S_{2n}(x) = \left(\frac{1}{1 + x^2} - \frac{1}{2 + x^2}\right) + \left(\frac{1}{3 + x^2} - \frac{1}{4 + x^2}\right) + \ldots + \left(\frac{1}{2n - 1 + x^2} - \frac{1}{2n + x^2}\right)$$

Since each bracket is positive, therefore $S_{2n}(x)$ is positive, increasing to its sum S(x).

$$\Rightarrow \quad S(x) - S_{2n}(x) > 0$$

Also,

$$S(x) - S_{2n}(x) = \frac{1}{2n + 1 + x^2} - \frac{1}{2n + 2 + x^2} + \frac{1}{2n + 3 + x^2} - \ldots$$

$$= \frac{1}{2n + 1 + x^2} - \left(\frac{1}{2n + 2 + x^2} - \frac{1}{2n + 3 + x^2} - \ldots\right)$$

$$< \frac{1}{2n + 1 + x^2} < \frac{1}{2n + 1}$$

$$\Rightarrow \qquad 0 < S(x) - S_{2n}(x) < \frac{1}{2n + 1} \qquad \ldots(1)$$

Again,

$$S_{2n+1}(x) - S(x) = \frac{1}{2n + 2 + x^2} - \frac{1}{2n + 3 + x^2} + \ldots$$

$$= \left(\frac{1}{2n + 2 + x^2} - \frac{1}{2n + 3 + x^2}\right) + \left(\frac{1}{2n + 4 + x^2} - \frac{1}{2n + 5 + x^2}\right) + \ldots$$

$$> 0$$

Also

$$S_{2n+1}(x) - S(x) = \frac{1}{2n+2+x^2} - \left(\frac{1}{2n+3+x^2} - \frac{1}{2n+4+x^2}\right) - \ldots$$

$$< \frac{1}{2n+2+x^2} < \frac{1}{2n+2}$$

$$\Rightarrow \qquad 0 < S_{2n+1}(x) - S(x) < \frac{1}{2n+2} \qquad \ldots(2)$$

Inequalities (1) and (2) imply that for any $\varepsilon > 0$, we can choose an integer m such that for all values of x,

$$|S(x) - S_n(x)| < \varepsilon, \quad \forall\, n \geq m$$

$\Rightarrow$ The series converges uniformly for all real values of x, and since each term of the series is continuous therefore the series will converge to a continuous sum function.

4.3 Uniform Convergence and Integration

Theorem 10 (A). *If a sequence $\{f_n\}$ converges uniformly to f on $[a, b]$, and each function f_n is integrable, then f is integrable on $[a, b]$, and the sequence $\left\{\int_a^x f_n\,dt\right\}$ converges uniformly to $\int_a^x f\,dt$ on $[a, b]$, i.e.,*

$$\int_a^x f\,dt = \lim_{n\to\infty} \int_a^x f_n\,dt, \quad \forall\, x \in [a, b] \qquad \ldots(1)$$

(B). *If a series Σf_n converges uniformly to f on $[a, b]$, and each term $f_n(x)$ is integrable, then f is integrable on $[a, b]$, and the series $\Sigma\left(\int_a^x f_n\,dt\right)$ converges uniformly to $\int_a^x f\,dt$ on $[a, b]$, i.e.,*

$$\int_a^x f\,dt = \sum_{n=1}^{\infty}\left(\int_a^x f_n\,dt\right), \quad \forall\, x \in [a, b] \qquad \ldots(2)$$

[We then say that the series is *integrable term by term*].

(A). Let $\varepsilon > 0$ be any number.

By the uniform convergence of the sequence, there exists an integer m such that for all $x \in [a, b]$

$$|f_n(x) - f(x)| < \frac{\varepsilon}{3(b-a)}, \quad \forall\, n \geq m$$

In particular,

$$|f_m(x) - f(x)| < \frac{\varepsilon}{3(b-a)} \qquad ...(3)$$

For this fixed m, since f_m is integrable, we choose a partition P of $[a, b]$, such that

$$U(P, f_m) - L(P, f_m) < \varepsilon/3 \qquad ...(4)$$

From equation (3),

$$f(x) < f_m(x) + \varepsilon/3(b-a)$$

$\Rightarrow$
$$U(P, f) < U(P, f_m) + \varepsilon/3 \qquad ...(5)$$

Again from equation (3),

$$f(x) > f_m(x) - \varepsilon/3(b-a)$$

$\Rightarrow$
$$L(P, f) > L(P, f_m) - \varepsilon/3 \qquad ...(6)$$

From equations (4), (5) and (6), we get

$$U(P, f) - L(P, f) < U(P, f_m) - L(P, f_m) + 2\varepsilon/3$$
$$< \varepsilon/3 + 2\varepsilon/3 = \varepsilon$$

$\Rightarrow$ f is integrable on $[a, b]$

We now proceed to prove relation (1).

Since the sequence $\{f_n\}$ converges uniformly to f, therefore for $\varepsilon > 0$, there exists an integer N such that for all $x \in [a, b]$,

$$|f_n(x) - f(x)| < \varepsilon/(b-a), \quad \forall\, n \geq N$$

Then for all $x \in [a, b]$ and for $n \geq N$, we have

$$\left| \int_a^x f\, dt - \int_a^x f_n\, dt \right| = \left| \int_a^x (f - f_n)\, dt \right| \leq \int_a^x |f - f_n|\, dt$$

$$< \frac{\varepsilon}{b-a}(x-a) \leq \varepsilon$$

$\Rightarrow$ $\left\{ \int_a^x f_n\, dt \right\}$ converges uniformly to $\int_a^x f\, dt$ over $[a, b]$, *i.e.*,

$$\int_a^x f\, dt = \lim_{n\to\infty} \int_a^x f_n\, dt, \quad \forall\, x \in [a, b]$$

(B). The proof may be supplemented by the reader himself.

Remark: The converse is neither asserted nor true, *i.e.*, a series (sequence) may converge to an integrable limit without being uniformly convergent. On the other hand, if the limit is not integrable or if integrable, the integral is not equal to the limit of the series (sequence) of integrals, *the convergence cannot be uniform.*

If the terms f_n are continuous for all n, a much shorter and simpler proof is possible.

Theorem 11. *If a series Σf_n uniformly converges to f on $[a, b]$ and each is f_n continuous on $[a, b]$, then f is integrable on $[a, b]$ and the series $\Sigma\left(\int_a^x f_n\, dt\right)$ converges uniformly to $\int_a^x f\, dt$, for all x in $[a, b]$, i.e.,*

$$\int_a^x f\, dt = \sum_{n=1}^{\infty}\left(\int_a^x f_n\, dt\right), \quad \forall\, x \in [a, b].$$

Since Σf_n is uniformly convergent to f on $[a, b]$ and each f_n is continuous on $[a, b]$, therefore the sum function f is continuous and hence integrable on $[a, b]$.

Again, since all the functions f_n are continuous, therefore the sum of a finite number of functions, $\sum_{r=1}^{n} f_r$ is also continuous and integrable on $[a, b]$, and

$$\sum_{r=1}^{n}\int_a^x f_r\, dt = \int_a^x \sum_{r=1}^{n} f_r\, dt$$

By the uniform convergence of the series, for $\varepsilon > 0$, we can find an integer N such that for all x in $[a, b]$,

$$\left|f - \sum_{r=1}^{n} f_r\right| < \varepsilon/(b-a), \quad \forall\, n \geq N$$

For such values of n, and all x in $[a, b]$

$$\left|\int_a^x f\, dt - \sum_{r=1}^{n}\int_a^x f_r\, dt\right| = \left|\int_a^x \left(f - \sum_{r=1}^{n} f_r\right) dt\right|$$

$$\leq \int_a^x \left|f - \sum_{r=1}^{n} f_r\right| dt$$

$$< \frac{\varepsilon}{b-a}\int_a^x dt \leq \varepsilon$$

$$\Rightarrow \quad \sum_{n=1}^{\infty}\left(\int_a^x f_n\, dt\right) \text{ converges uniformly to } \int_a^x f\, dt \text{ on } [a, b],$$

i.e.,
$$\int_a^x f\, dt = \sum_{n=1}^{\infty}\int_a^x f_n\, dt, \text{ for all } x \in [a, b].$$

The corresponding theorem for sequences may be stated and proved in the same way.

4.4 Uniform Convergence and Differentiation

The series $\sum \frac{\sin nx}{n^2}$ converges uniformly for all values of x and for every term, without exception, is continuous and differentiable. The series of differentials $\sum \frac{\cos nx}{n}$, however, diverges at $x = 0$. The situation, therefore, seems to be different in the case of differentiation, and accordingly the theorem on term-by-term differentiation must be of a different stamp.

Theorem 12 (A). *Let $\{f_n\}$ be a sequence of differentiable functions on $[a, b]$ such that it converges at least at one point $x_0 \in [a, b]$. If the sequence of differentials $\{f_n'\}$ converges uniformly to G on $[a, b]$, then the given sequence $\{f_n\}$ converges uniformly on $[a, b]$ to f and $f'(x) = G(x)$.*

(B). *Let $\sum f_n$ be a series of differentiable functions on $[a, b]$ such that it converges at least at one point $x_0 \in [a, b]$. If the series of differentials $\sum f_n'$ converges uniformly to G on $[a, b]$, then the given series $\sum f_n$ converges uniformly on $[a, b]$ to f, and $f'(x) = G(x)$.*

[The existence of f' is a part of the conclusion.]

(A). Let $\varepsilon > 0$ be any number.

By the convergence of $\{f_n(x_0)\}$ and uniform convergence of $\{f_n'\}$, for $\varepsilon > 0$, we can choose a positive integer N such that for all $x \in [a, b]$,

$$|f_{n+p}(x_0) - f_n(x_0)| < \varepsilon/2, \qquad \forall\, n \geq N,\ p \geq 1 \qquad ...(1)$$

$$|f_{n+p}'(x) - f_n'(x)| < \varepsilon/2(b-a), \quad \forall\, n \geq N,\ p \geq 1 \qquad ...(2)$$

Applying Lagrange's mean value theorem to the function $(f_{n+p} - f_n)$ for any two points x and t of $[a, b]$, we get for $x < \xi < t$, for all $n \geq N,\ p \geq 1$

$$|f_{n+p}(x) - f_n(x) - f_{n+p}(t) + f_n(t)| = |x - t|\,|f_{n+p}'(\xi) - f_n'(\xi)|$$

$$< \frac{|x-t|\,\varepsilon}{2(b-a)} \qquad ...(3)$$

$$< \varepsilon/2 \qquad ...(3A)$$

and

$$|f_{n+p}(x) - f_n(x)| \le |f_{n+p}(x) - f_n(x) - f_{n+p}(x_0) + f_n(x_0)| + |f_{n+p}(x_0) - f_n(x_0)|$$

$$< \frac{1}{2}\varepsilon + \frac{1}{2}\varepsilon = \varepsilon \qquad \text{[using (1) \& (3A)]}$$

$\Rightarrow$ The sequence $\{f_n\}$ uniformly converges on $[a, b]$.

Let it converges to f, say.

For a fixed x on $[a, b]$ and for $t \in [a, b],\ t \neq x,$ let us define

$$\phi_n(t) = \frac{f_n(t) - f_n(x)}{t - x},\ n = 1,2,3,\ldots \qquad \text{...(4)}$$

$$\phi(t) = \frac{f(t) - f(x)}{t - x} \qquad \text{...(5)}$$

Since each f_n is differentiable, therefore for each n

$$\lim_{t \to x} \phi_n(t) = f_n'(x) \qquad \text{...(6)}$$

$\therefore$

$$|\phi_{n+p}(t) - \phi_n(t)| = \frac{1}{|t - x|}|f_{n+p}(t) - f_{n+p}(x) - f_n(t) + f_n(x)|$$

$$= \frac{1}{|t - x|}\,|\,\{f_{n+p}(t) - f_n(t)\} - \{f_{n+p}(x) - f_n(x)\}\,|$$

$$< \frac{\varepsilon}{2(b - a)},\quad \forall\ n \ge N,\ p \ge 1 \qquad \text{[using (3)]}$$

so that $\{\phi_n(t)\}$ converges uniformly on $[a, b]$, for $t \neq x$.

Since $\{f_n\}$ also converges uniformly on f, therefore from (4),

$$\lim_{n\to\infty} \phi_n(t) = \lim_{n\to\infty} \frac{f_n(t) - f_n(x)}{t - x} = \frac{f(t) - f(x)}{t - x} = \phi(t)$$

Thus $\{\phi_n(t)\}$ converges uniformly to $\phi(t)$ on $[a, b]$, for $t \in [a, b],\ t \neq x$.

Applying Theorem 7(A) to the uniformly convergent sequence $\{\phi_n(t)\}$ and using (6), we get

$$\lim_{t \to x} \phi(t) = \lim_{n\to\infty} f_n'(x) = G(x)$$

$$\Rightarrow \qquad \lim_{t \to x} \phi(t) \text{ exists,}$$

and therefore (5) implies that f is differentiable and

$$\lim_{t \to x} \phi(t) = f'(x)$$

Hence,

$$f'(x) = G(x) = \lim_{n\to\infty} f_n'(x)$$

This completes the proof of the theorem.

If in addition to the above hypothesis of the theorem, continuity of the functions f_n' is also assumed, then a much shorter proof of the theorem exists.

We now prove a simple version of the theorem for series, that for sequence may be written down on the same lines.

Theorem 13. *Let a series Σf_n of differentiable functions converges pointwise to f on [a, b] and each f_n' is continuous on [a, b], and the series $\Sigma f_n'$ converges uniformly to G on [a, b], then the given series Σf_n converges uniformly to f on [a, b], and $f'(x) = G(x)$.*

Since the series $\Sigma f_n'$ of continuous functions converges uniformly to G on $[a, b]$, therefore its sum function G is continuous on $[a, b]$, and consequently the function

$$\int_a^x G(t)\, dt \text{ is differentiable, and}$$

$$\frac{d}{dx}\int_a^x G(t)\, dt = G(x), \text{ for all } x \in [a, b]. \qquad \text{...(1)}$$

For every $x \in [a, b]$, $f(x) = \sum_{n=1}^{\infty} f_n(x)$.

Now, since each function f_n', being continuous, is integrable on $[a, b]$, and so by the fundamental theorem of calculus,

$$\int_a^x f_n'(t)\, dt = f_n(x) - f_n(a), \text{ for all } n \geq 1, x \in [a, b]$$

$$\therefore \qquad \sum_{n=1}^{\infty}\int_a^x f_n'(t)\, dt = f(x) - f(a), \text{ for all } x \in [a, b] \qquad \text{...(2)}$$

Again, since the series $\Sigma f_n'$, of integrable functions, converges uniformly to G on $[a, b]$, therefore term-by-term integration is valid, *i.e.*,

$$\int_a^x G(t)\, dt = \sum_{n=1}^{\infty}\int_a^x f_n'(t)\, dt, \ \forall\ x \in [a, b] \qquad \text{...(3)}$$

From equations (1), (2) and (3), it follows that

$$f'(x) = G(x), \text{ for all } x \in [a, b].$$

or equivalently

$$\frac{d}{dx}\sum_{n=1}^{\infty} f_n(x) = \sum_{n=1}^{\infty}\frac{d}{dx} f_n(x), a \leq x \leq b$$

i.e., the term-by-term differentiation of the series is valid.

4.5 Some Associated Examples

We now consider some examples to show that the limit functions of uniformly convergent series and sequences of continuous (integrable) functions are continuous (integrable) but there do exist series and sequences which though not uniformly convergent but still possess continuous (integrable) functions.

Example 19. The series

$$1 - x + x^2 - x^3 + \text{K} = \frac{1}{1+x}, \quad 0 < x < 1$$

- Each term is integrable.

  Integrating from 0 to 1, the right hand side gives

$$\int_0^1 \frac{dx}{1+x} = \log 2$$

while the other side gives

$$\left[x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \ldots\right]_0^1 = 1 - \frac{1}{2} + \frac{1}{2} - \frac{1}{2} + \ldots$$

But we know that

$$\log 2 = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \ldots$$

Thus the two sides are equal at $x = 1$, and so term by term integration is possible over [0, 1], even though the given series is not uniformly convergent on [0, 1].

Example 20. The sequence $\{f_n\}$, where

$$f_n(x) = nxe^{-nx^2}, \; n = 1, 2, 3, \ldots$$

converges pointwise to zero on [0, 1].

- Here

$$\int_0^1 f \; dx = 0$$

and

$$\int_0^1 f_n dx = \frac{1}{2}\left[-e^{-nx^2}\right]_0^1 = \frac{1}{2}(1 - e^{-n})$$

$$\therefore \quad \lim_{n\to\infty} \int_0^1 f_n d_n = \lim_{n\to\infty} \frac{1}{2}\left(1 - e^{-n}\right) = \frac{1}{2} \neq \int_0^1 f \; dx$$

$\Rightarrow$ convergence cannot be uniform on [0, 1].

Note: If we, first, show that the sequence is non-uniformly convergent, then this is an example of a sequence which, though not uniformly convergent, yet has an integrable limit function.

Ex. Show although $f_n(x) = x \exp(-nx^2)$ is uniformly convergent in $[-1, 1]$ to a differentiable function, the limit and differentiation process cannot be interchanged.

Example 21. Show that for $-1 < x < 1$,

$$\frac{1}{1+x} + \frac{2x}{1+x^2} + \frac{4x^3}{1+x^4} + \frac{8x^7}{1+x^8} + \ldots = \frac{1}{1-x}$$

■ Consider the series

$$\log(1-x) + \log(1+x) + \log(1+x^2) + \log(1+x^4) + \ldots \qquad \ldots(1)$$

The nth partial sum

$$S_n = \log\{(1-x)(1+x)(1+x^2)(1+x^4)\ldots(1+x^{2n-1})\}$$

$$= \log(1-x^{2n}) \to 0 \text{ as } n \to \infty \text{ for } |x| < 1$$

Hence series (1) converges to zero.

The series of differentials of (1), ignoring the first two terms, is

$$\frac{2x}{1+x^2} + \frac{4x^3}{1+x^4} + \frac{8x^7}{1+x^8} + \ldots + \frac{2^n x^{2n-1}}{1+x^{2n}} + \ldots \qquad \ldots(2)$$

Now

$$\left|\frac{2^n x^{2n-1}}{1+x^{2n}}\right| \leq 2^n \rho^{2n-1}, \text{ for } |x| \leq \rho < 1$$

The series $\Sigma\, 2^n \rho^{2n-1}$ is convergent and therefore by M-test, the series of differentials (2) converges uniformly for $|x| \leq \rho < 1$, and therefore (by Theorem 12) its sum is the differential of the sum of series (1) without the first two terms.

But the sum of series (1) without the first two terms $= -\log(1-x) - \log(1+x)$.

Hence,

$$\frac{2x}{1+x^2} + \frac{4x^3}{1+x^4} + \frac{8x^7}{1+x^8} + \ldots = \frac{d}{dx}\{-\log(1-x) - \log(1+x)\}$$

$$= \frac{1}{1-x} - \frac{1}{1+x}$$

or

$$\frac{1}{1+x} + \frac{2x}{1+x^2} + \frac{4x^3}{1+x^4} + \ldots = \frac{1}{1-x}$$

Example 22. The sequence $\{f_n\}$, where $f_n(x) = \dfrac{nx}{1+n^2x^2}$ converges to f and $f(x) = 0$, for all real x.

■ Clearly $f'(x) = 0$

and
$$f_n'(x) = \frac{n(1-n^2x^2)}{(1+n^2x^2)^2}$$

Therefore, when $x \neq 0$, $f_n'(x) \to 0$, as $n \to \infty$ and so the formula

$$f'(x) = \lim_{n\to\infty} f_n'(x), \text{ is true.}$$

But at $x = 0$,

$$f_n'(0) = \lim_{x\to 0} \frac{n}{1+n^2x^2} = n, \text{ which tends to } \infty, \text{ as } n \to \infty$$

Thus at $x = 0$, $f'(x) = \lim_{n\to\infty} f_n'(x)$ is false.

So here it is $\{f_n'\}$ that does not converge uniformly in an interval that contains zero.

Example 23. Show that the sequence $\{f_n\}$, where

$$f_n(x) = \frac{x}{1+nx^2}$$

converges uniformly to a function f on $[0, 1]$, and that the equation

$$f'(x) = \lim_{n\to\infty} f_n'(x)$$

is true if $x \neq 0$ and false if $x = 0$. Why so?

■ It may be easily shown (Example 7) that the sequence $\{f_n\}$ converges uniformly to zero for all real x.

The limit function $f(x) = 0$.

When $x \neq 0$,

$$f_n'(x) = \frac{1-nx^2}{(1+nx^2)^2} \to 0, \text{ as } n \to \infty$$

so that if $x \neq 0$ the formula $f'(x) = \lim_{n\to\infty} f_n'(x)$, is true.

At $x = 0$,

$$f_n'(0) = \lim_{x\to 0} \frac{1}{1+nx^2} = 1$$

so that

$$\lim_{n\to\infty} f_n'(0) = 1 \neq f'(0)$$

Hence at $x = 0$, the formula $f'(x) = \lim_{n\to\infty} f_n'(x)$, is false.

It is so because, the sequence $\{f_n'\}$ is not uniformly convergent in any interval containing zero.

Example 24. Show that the sequence $\{f_n\}$, where

$$f_n(x) = \frac{\log(1 + n^3x^2)}{n^2}$$

is uniformly convergent on the interval [0, 1].

■ The sequence $\{\phi_n\}$, where $\phi_n(x) = \dfrac{2nx}{1 + n^3x^2} \equiv f_n'(x)$, may be easily shown to be uniformly convergent to ϕ, where $\phi(x) = 0$ on [0, 1]. Also each function ϕ_n is continuous on the given interval.

Therefore, (by Theorem 10) the sequence of its integrals, $\{f_n\}$ converges uniformly to $\int_0^x \phi\, dt = 0$ on [0, 1].

Hence the result.

Example 25. Show that the sequence $\{f_n\}$, where

$$f_n(x) = \begin{cases} n^2x, & 0 \le x \le 1/n \\ -n^2x + 2n, & 1/n \le x \le 2/n \\ 0, & 2/n \le x \le 1 \end{cases}$$

is not uniformly convergent on [0, 1].

■ The sequence converges to f, where $f(x) = 0$, for all $x \in [0, 1]$. Each function f_n and f are continuous on [0, 1].

Also $$\int_0^1 f_n\, dx = \int_0^{1/n} n^2x\, dx + \int_{1/n}^{2/n} (-n^2x + 2n)\, dx + \int_{2/n}^1 0\, dx = 1$$

But $$\int_0^1 f\, dx = 0$$

$\therefore$ $$\lim_{n\to\infty} \int_0^1 f_n\, dx \ne \int_0^1 f\, dx$$

So (by Theorem 10) the sequence $\{f_n\}$ cannot converge uniformly on [0, 1].

EXERCISE

1. Show that the following series converge uniformly:

   (i) $e^x + e^{2x} + e^{3x} + \ldots, |x| \le \dfrac{1}{4}$.

CHAPTER 12

(*ii*) $x - x^2 + x^3 - x^4 + \ldots, -\frac{1}{2} \le x \le \frac{1}{2}.$

2. Discuss the uniform convergence of

$$1 + \frac{e^{-2x}}{2^2 - 1} - \frac{e^{-4x}}{4^2 - 1} + \frac{e^{-6x}}{6^2 - 1} - \ldots, \text{ for all real } x \ge 0$$

3. Show that the sequence $\{f_n\}$, where $f_n(x) = x - \frac{x^n}{n}$, converges uniformly on [0, 1]. Show also that the sequence $\{f_n'\}$ of differentials does not converge uniformly on [0, 1].

4. Decide whether or not the sequences $\{f_n'\}$ and $\left\{\int_0^x f_n \, dt\right\}$ converge uniformly on [0, 1], where

(*i*) $f_n(x) = \frac{x}{1 + n^2 x}$, (*ii*) $f_n(x) = nxe^{-nx^2}$,

(*iii*) $f_n(x) = \frac{n^2 x}{1 + n^3 x^2}$, (*iv*) $f_n(x) = \frac{2 + nx^2}{2 + nx}$.

5. Show that the sequence $\{(\sin x)^{1/n}\}$ converges but not uniformly on $[0, \pi]$.

6. Show that the sequence $\left\{\left(\frac{\sin x}{x}\right)^{1/n}\right\}$ converges but not uniformly on $[0, \pi]$.

7. Show that the series $\sum_1^{\infty} f_n$, where $f_n(x) = \frac{x^2}{(1 + x^2)^n}$, does not converge uniformly for $x \ge 0$.

[*Hint:* Each f_n is continuous but the series converges to a discontinuous function].

8. Show that the sequence $\{f_n\}$, where $f_n(x) = nx\,(1 - x^2)^n$, converges but not uniformly to f, where $f(x) = 0$, for $0 \le x \le 1$, and that

$$\lim_{n \to \infty} \int_0^1 f_n \, dx \ne \int_0^1 f \, dx.$$

9. Show that the sequence $\{f_n\}$ defined on [0, 1] by

$$f_n(x) = \begin{cases} n(1 - nx), & 0 < x < 1/n \\ 0, & \text{otherwise} \end{cases}$$

converges pointwise, but not uniformly to f, where $f(x) = 0$ for $0 \le x \le 1$, and that

$$\lim_{n \to \infty} \int_0^1 f_n \, dx \ne \int_0^1 f \, dx.$$

10. Consider the sequence $\{f_n\}$ of functions defined by

(*i*) $f_n(x) = |x|^{1+1/n}$, $x \in [-1, 1]$,

(*ii*) $f_n(x) = (2x/\pi)\tan^{-n}(nx)$, $x \in \mathbf{R}$.

Show that $\{f_n\}$ converges uniformly to $|x|$, but $\{f_n'\}$ converges only pointwise to sgn (x), on the indicated interval.

11. Given $f(x) = \sum_{n=1}^{\infty} \frac{1}{n^3 + n^4 x^2}$, justify the validity of the equation

$$f'(x) = -2x \sum_{n=1}^{\infty} \frac{1}{n^2 (1 + nx^2)^2}.$$

12. Show that, if a sequence of $\{f_n\}$ bounded functions on $[a, b]$ converges uniformly to f on $[a, b]$, then f is bounded on $[a, b]$. Is the result still valid, if we have only pointwise convergence?

13. The sequence $\{f_n\}$ is defined for $x \geq 0$, as follows:

$$f_1(x) = \sqrt{x},\ f_{n+1}(x) = \sqrt{x + f_n(x)},\ n \geq 1.$$

Show that $\{f_n\}$ converges uniformly on $[a, b]$, $0 < a < b$. Is the convergence uniform on $[0, 1]$?

5. THE WEIERSTRASS APPROXIMATION THEOREM

We shall now study a very famous theorem, discovered originally by Weierstrass. The theorem is described by saying that every continuous function can be 'uniformly approximated' by polynomials to within any degree of accuracy. Many proofs of this classical theorem are known, and the one we give is perhaps as concise and instructive as most.

Theorem 14. *If f is a real continuous function defined on a closed interval $[a, b]$ then there exists a sequence of real polynomials $\{P_n\}$ which converges uniformly to $f(x)$ on $[a, b]$, i.e, $\lim_{n\to\infty} P_n(x) = f(x)$, converges uniformly on $[a, b]$.*

If $a = b$, the conclusion follows by taking $P_n(x)$ to be a constant polynomial, defined by $P_n(x) = f(a)$, for all n.

We may thus assume that $a < b$.

We next observe that a linear transformation $x' = (x - a)/(b - a)$ is a continuous mapping of $[a, b]$ onto $[0, 1]$. Accordingly, we assume without loss of generality that $a = 0$, $b = 1$.

Consider

$$F(x) = f(x) - f(0) - x\,[f(1) - f(0)], \quad \text{for } 0 \leq x \leq 1$$

Here $F(0) = 0 = F(1)$, and if F can be expressed as a limit of a uniformly convergent sequence of polynomials, then the same is true for f, since $f - F$ is a polynomial. So we may assume that $f(1) = f(0) = 0$.

Let us further define $f(x)$ to be zero for x outside $[0, 1]$. Thus, f is now uniformly continuous on the whole real line.

Let us consider the polynomial (non-negative for $|x| \leq 1$).

$$B_n(x) = C_n(1 - x^2)^n, \quad n = 1, 2, 3, \ldots \qquad \text{...(1)}$$

where C_n, independent of x, is so chosen that

$$\int_{-1}^{1} B_n(x)\,dx = 1 \quad \text{for } n = 1, 2, 3, \ldots \qquad \text{...(2)}$$

$$\therefore \qquad 1 = \int_{-1}^{1} C_n(1-x^2)^n\, dx = 2C_n \int_0^1 (1-x^2)^n\, dx$$

$$\geq 2C_n \int_0^{1/\sqrt{n}} (1-x^2)^n\, dx$$

$$\geq 2C_n \int_0^{1/\sqrt{n}} (1-nx^2)\, dx = \frac{4C_n}{3\sqrt{n}} > \frac{C_n}{\sqrt{n}}$$

$$\Rightarrow \qquad C_n < \sqrt{n} \qquad \text{...(3)}$$

which gives some information about the order of magnitude of C_n.

Therefore, for *any* $\delta > 0$, (3) gives

$$B_n(x) \leq \sqrt{n}(1-\delta^2)^n, \text{ when } \delta \leq |x| \leq 1 \qquad \text{...(4)}$$

so that $B_n \to 0$ uniformly, for $\delta \leq |x| \leq 1$.

Again, let

$$P_n(x) = \int_{-1}^{1} f(x+t)\, B_n(t)\, dt, \quad 0 \leq x \leq 1 \qquad \text{...(5)}$$

$$= \int_{-1}^{-x} f(x+t)\, B_n(t)\, dt + \int_{-x}^{1-x} f(x+t)\, B_n(t)\, dt + \int_{1-x}^{1} f(x+t)\, B_n(t)\, dt$$

For $|x| \leq 1, -1+x \leq x+t \leq 0$, for $-1 \leq t \leq -x$, so that $x+t$ lies outside [0, 1] and therefore $f(x+t) = 0$, and hence the first integral on the right vanishes. Similarly, the third integral is also equal to zero. Hence

$$P_n(x) = \int_{-x}^{1-x} f(x+t)\, B_n(t)\, dt = \int_0^1 f(t)\, B_n(t-x)\, dt$$

which clearly is a real polynomial.

We now proceed to show that the sequence $\{P_n(x)\}$ converges uniformly to f on [0, 1].

Continuity of f on the closed interval [0, 1] implies that f is bounded and uniformly continuous on [0, 1].

Therefore, there exists M such that

$$M = \sup |f(x)| \qquad \text{...(6)}$$

and for any $\varepsilon > 0$, we can choose $\delta > 0$ such that for any two points x_1, x_2 in [0, 1],

$$|f(x_1) - f(x_2)| < \varepsilon/2, \text{ when } |x_1 - x_2| < \delta \leq 1 \qquad \text{...(7)}$$

For $0 \leq x \leq 1$, we have

$$|P_n(x) - f(x) = \left| \int_{-1}^{1} f(x+t)\, B_n(t)\, dt - f(x) \right|$$

$$= \left| \int_{-1}^{1} \{f(x+t) - f(x)\}\, B_n(t)\, dt \right| \qquad \text{[using (2)]}$$

$$\leq \int_{-1}^{1} | f(x+t) - f(x) | B_n(t)\, dt \qquad (\because\ B_n(t) \geq 0)$$

$$= \int_{-1}^{-\delta} |f(x+t) - f(x)| B_n(t)\, dt + \int_{-\delta}^{\delta} |f(x+t) - f(x)| B_n(t)\, dt$$

$$+ \int_{\delta}^{1} |f(x+t) - f(x)| B_n(t)\, dt$$

$$\leq 2M \int_{-1}^{-\delta} B_n(t)\, dt + \frac{\varepsilon}{2} \int_{-\delta}^{\delta} B_n(t)\, dt + 2M \int_{\delta}^{1} B_n(t)\, dt$$

[using equations (6) and (7)]

$$\leq 2M\sqrt{n}\,(1-\delta^2)^n \left\{ \int_{-1}^{-\delta} dt + \int_{\delta}^{1} dt \right\} + \varepsilon/2$$

[using equations (2) and (4)]

$$\leq 4M\sqrt{n}(1-\delta^2)^n + \varepsilon/2$$

$< \varepsilon$, for large values of n.

Thus for $\varepsilon > 0$, there exists N (independent of x) such that

$$| P_n(x) - f(x) | < \varepsilon, \quad \forall\ n \geq N$$

$$\Rightarrow \qquad \lim_{n\to\infty} P_n(x) = f(x), \text{ uniformly on } [0, 1].$$

Second method. If $a = b$, the conclusion follows by taking $P_n(x)$ to be a constant polynomial, defined by $P_n(x) = f(a)$, for all n.

We may thus assume that $a < b$.

We next observe that a linear transformation $x' = (x-a)/(b-a)$ is a continuous mapping of $[a, b]$ onto $[0, 1]$. Accordingly, we assume without loss of generality that $a = 0$, $b = 1$.

For positive integers n and k when $0 \leq k \leq n$, the binomial coefficient $\binom{n}{k}$ is defined by

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

Let us define the polynomials B_n, where

$$B_n(x) = \sum_{k=0}^{n} \binom{n}{k} x^k (1-x)^{n-k}\ f(k/n),\ n = 1, 2, 3, \ldots, \text{ and } x \in [0, 1]$$

called the *Bernstein polynomials* associated with f.

Let us first consider some identities which will be our main tools to show that certain Bernstein polynomials exist which continuously converge to f on $[0, 1]$.

The first of the identities is a special case of the binomial theorem,

$$\sum_{k=0}^{n} \binom{n}{k} x^k (1-x)^{n-k} = [x + (1-x)]^n = 1 \qquad ...(1)$$

Differentiating with respect to x, we get

$$\sum_{k=0}^{n} \binom{n}{k} [kx^{k-1}(1-x)^{n-k} - (n-k)\, x^k (1-x)^{n-k-1}] = 0$$

or
$$\sum_{k=0}^{n} \binom{n}{k} x^{k-1}(1-x)^{n-k-1}\,(k-nx) = 0$$

and multiplication by $x(1-x)$ gives

$$\sum_{k=0}^{n} \binom{n}{k} x^k (1-x)^{n-k}\,(k-nx) = 0 \qquad ...(2)$$

Differentiating again with respect to x, we get

$$\sum_{k=0}^{n} \binom{n}{k} [-nx^k(1-x)^{n-k} + x^{k-1}(1-x)^{n-k-1}(k-nx)^2] = 0$$

which on applying in (1), gives

$$\sum_{k=0}^{n} \binom{n}{k} x^{k-1}(1-x)^{n-k-1}\,(k-nx)^2 = n$$

and on multiplying by $x(1-x)$, we get

$$\sum_{k=0}^{n} \binom{n}{k} x^k (1-x)^{n-k}\,(k-nx)^2 = nx(1-x)$$

or
$$\sum_{k=0}^{n} \binom{n}{k} x^k (1-x)^{n-k}\,(x-k/n)^2 = \frac{x(1-x)}{n} \qquad ...(3)$$

The maximum value of $x(1-x)$ in $[0, 1]$ being $\frac{1}{4}$.

$\therefore$
$$\sum_{k=0}^{n} \binom{n}{k} x^k (1-x)^{n-k}\,(x-k/n)^2 \le \frac{1}{4n} \qquad ...(4)$$

Continuity of f on the closed interval $[0, 1]$ implies that f is bounded and uniformly continuous on $[0, 1]$.

Hence, there exists $K > 0$, such that

$$|f(x)| \le K, \quad \forall\, x \in [0, 1]$$

and for any $\varepsilon > 0$, there exists $\delta > 0$ such that for all $x \in [0, 1]$

$$|f(x) - f(k/n)| < \frac{1}{2}\varepsilon, \text{ when } |x - k/n| < \delta \qquad ...(5)$$

For any fixed but arbitrary x in [0, 1], the values 0, 1, 2, 3, ..., n of k may be divided into two parts:

Let A be the set of values of k for which $|x - k/n| < \delta$, and B the set of the remaining values, for which $|x - k/n| \geq \delta$.

For $k \in B$, using (4),

$$\sum_{k \in B} \binom{n}{k} x^k (1-x)^{n-k} \delta^2 \leq \sum_{k \in B} \binom{n}{k} x^k (1-x)^{n-k} (x - k/n)^2 \leq \frac{1}{4n}$$

$$\Rightarrow \qquad \sum_{k \in B} \binom{n}{k} x^k (1-x)^{n-k} \leq \frac{1}{4n\delta^2} \qquad ...(6)$$

Using (1), we see that for this fixed x in [0, 1],

$$\begin{aligned} |f(x) - B_n(x)| &= \left| \sum_{k=0}^{n} \binom{n}{k} x^k (1-x)^{n-k} [f(x) - f(k/n)] \right| \\ &\leq \sum_{k=0}^{n} \binom{n}{k} x^k (1-x)^{n-k} \, |f(x) - f(k/n)| \end{aligned}$$

Thus, summation on the right may be split into two parts, according as $|x - k/n| < \delta$ or $|x - k/n| \geq \delta$. Thus

$$\begin{aligned} |f(x) - B_n(x)| &\leq \sum_{k \in A} \binom{n}{k} x^k (1-x)^{n-k} \, |f(x) - f(k/n)| \\ &\quad + \sum_{k \in B} \binom{n}{k} x^k (1-x)^{n-k} \, |f(x) - f(k/n)| \\ &< \frac{\varepsilon}{2} \sum_{k \in A} \binom{n}{k} x^k (1-x)^{n-k} + 2K \sum_{k \in B} \binom{n}{k} x^k (1-x)^{n-k} \\ &\leq \varepsilon/2 + 2K/4n\delta^2 < \varepsilon, \text{ using equations (1), (5) and (6),} \end{aligned}$$

for values of n greater than $K/\varepsilon\delta^2$.

Thus $\{B_n(x)\}$ converges uniformly to $f(x)$ on [0, 1].

Corollary. *For any interval* $[-a, a]$ *there is a sequence of real polynomials* P_n *such that* $P_n(0) = 0$, *and that*

$$\lim_{n \to \infty} P_n(x) = |x|, \text{ uniformly on } [-a, a].$$

Since $|x|$ is a real continuous function on $[-a, a]$, therefore by 'Weierstrass approximation theorem' there exists a sequence $\{Q_n\}$ of real polynomials which converges uniformly to $|x|$ on $[-a, a]$.

In particular,

$$Q_n(0) \to 0, \text{ as } n \to \infty$$

Hence, the polynomials

$$P_n(x) - Q_n(x) - Q_n(0), \; n = 1, 2, 3, \ldots$$

have the required property.

Example 26. If f is continuous on $[0, 1]$, and if

$$\int_0^1 x^n f(x)\, dx = 0, \text{ for } n = 0, 1, 2, \ldots \qquad \ldots(1)$$

then show that $f(x) = 0$ on $[0, 1]$.

■ From (1), it follows that the integral of the product of f with any polynomial is zero.

Now, since f is continuous on $[0, 1]$, therefore, by 'Weierstrass approximation theorem', there exists a sequence $\{P_n\}$ of polynomials, such that $P_n \to f$ uniformly on $[0, 1]$. And so $P_n f \to f^2$ uniformly on $[0, 1]$. Since f, being continuous, is bounded on $[0, 1]$. Therefore, by Theorem 10 (A),

$$\int_0^1 f^2\, dx = \lim_{n\to\infty} \int_0^1 P_n f\, dx = 0 \qquad \text{[using (1)]}$$

$$\therefore \qquad f^2 = 0 \text{ on } [0, 1]. \quad \text{Hence, } f = 0 \text{ on } [0, 1].$$

Ex. 1. Obtain the first six Bernstein polynomials for the following functions:

(*i*) $f(x) = |x|, [-1, 1]$ (*ii*) $f(x) = \sin x, [0, \pi]$

(*iii*) $f(x) = e^x, [-1, 1]$ (*iv*) $f(x) = \cos 6x, [0, \pi]$.

Ex. 2. Show that there does not exist a sequence of polynomials converging uniformly on **R** to f, where

(*i*) $f(x) = e^x$ (*ii*) $f(x) = \sin x$.

Ex. 3. Show that, if f is continuous on **R**, then there exists a sequence $\{P_n\}$ of polynomials converging uniformly to f on each bounded subset of **R**.

[*Hint*: Arrange for $|P_n(x) - f(x)| < 1/n$, for $|x| \le n$].

13 Power Series

1. INTRODUCTION

The terms of the series which we have examined so far (with the exception of those considered in the chapter on Uniform Convergence) were for the most part, determinate numbers. In such cases the series may be characterised at having constant terms. This, however, was not everywhere the case. In the geometric series Σr^n, for instance, the terms only become determinate when the value of r is assigned. Our investigation of the behaviour of this series did not terminate with the mere statement of the convergence or divergence, the result was: the series converges if $|r| < 1$, but diverges if $|r| \geq 1$. The solution thus depends, as do the terms of the series, on the value of a quantity left undetermined—a *variable*. In this chapter we propose only to consider, in detail within the scope of the present work, series whose *generic term* has the form $a_n x^n$, *i.e.*, we shall consider series of the form

$$a_0 + a_1 x + a_2 x^2 + ... + a_n x^n + ... \equiv \sum_{n=0}^{\infty} a_n x^n$$

Such series are called *power series* (in x) and the numbers a_n (dependent on n but not on x) their *coefficients*.

2. DEFINITIONS

For $x = 0$, obviously *every* power series is convergent whatever be the value of the coefficients. The most important fact about a power series is that either:

(*i*) it converges for no value of x other than the self-evident point $x = 0$, we then say that it is *nowhere convergent*, *e.g.*, $\Sigma n^n x^n$ converges for no value of x other than $x = 0$,

or

(*ii*) it converges for all values of x, and is then called *everywhere convergent*,

e.g., $$\Sigma \frac{x^n}{n!}, \Sigma (-1)^n \frac{x^n}{n!},$$

or

(*iii*) (*the general case*) it converges for some values of x and diverges for others—the totality of points x for which it converges is called its *region of convergence*.

Thus, if $\Sigma a_n x^n$ is a power series which does not converge everywhere or nowhere, then a definite positive number R exists such that $\Sigma a_n x^n$ converges (indeed absolutely) for every $|x| < R$ but diverges

for every $|x| > R$. The number R, which is associated with every power series, is called the *radius of convergence* and the interval, $]-R, R[$, the *interval of convergence*, of the given power series.

The behaviour of a power series at $|x| = R$, depends entirely upon the character of the sequence $\{a_n\}$ of its coefficients. For instance, both the series

$$\Sigma \frac{x^n}{n^2}, \Sigma \frac{x^n}{n}$$

converge when $|x| < 1$ and diverge when $|x| > 1$. When $|x| = 1$, the first series converges while the second diverges at $x = 1$, and converges at $x = -1$.

Theorem 1. *If $\overline{\lim} |a_n|^{1/n} = \frac{1}{R}$, then the series $\Sigma a_n x^n$ is convergent (in fact absolutely) for $|x| < R$ and divergent for $|x| > R$.*

Now
$$\overline{\lim_{x\to\infty}} \left| a_n x^n \right|^{1/n} = \frac{|x|}{R}$$

Hence by Cauchy's root test, the series $\Sigma a_n x^n$ is absolutely convergent and therefore convergent for $|x| < R$ and divergent for $|x| > R$.

Definition. In view of the above theorem, the radius of convergence R of a power series is defined to be equal to

$$\frac{1}{\overline{\lim} |a_n|^{1/n}}, \text{ when } \overline{\lim} |a_n|^{1/n} > 0$$

$$\infty \qquad \text{when } \overline{\lim} |a_n|^{1/n} = 0$$

$$0 \qquad \text{when } \overline{\lim} |a_n|^{1/n} = \infty$$

Thus for a nowhere convergent power series, $R = 0$, while for an everywhere convergent power series, $R = \infty$.

ILLUSTRATIONS

1. The series $1 + 2x + 3x^2 + 4x^3 + \ldots$, has radius of convergence = 1.
2. The series $\sum_{1}^{\infty} \frac{x^{n-1}}{n^2}$ converges absolutely for $x \in]-1, 1[$.
3. The radius of convergence of the series $1 - \frac{x}{2} + \frac{x^2}{3} - \frac{x^3}{4} + \ldots$ is 1. It is also convergent at $x = 1$.
4. The series $1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \ldots$, has infinite radius of convergence. It converges (absolutely) for all values of x.
5. The series $1 + x + 2!x^2 + 3!x^3 + \ldots$, does not converge for any value of x (other than 0). Its radius of convergence $R = 0$.

2.1 Some Theorems

Theorem 2. *If a power series $\Sigma a_n x^n$ converges for $x = x_0$ then it is absolutely convergent for every $x = x_1$, when $|x_1| < |x_0|$.*

Since the series $\Sigma a_n x_0^n$ is convergent, therefore $a_n x_0^n \to 0$, as $n \to \infty$.

Thus, for $\varepsilon = \frac{1}{2}$ (say), there exists an integer N such that

$$\left|a_n x_0^n\right| < \frac{1}{2}, \text{ for } n \geq N, \text{ and so}$$

$$\left|a_n x_1^n\right| = \left|a_n x_0^n\right| \cdot \left|\frac{x_1}{x_0}\right|^n < \frac{1}{2}\left|\frac{x_1}{x_0}\right|^n, \text{ for } n \geq N$$

But $\Sigma\left|\frac{x_1}{x_0}\right|^n$ is a convergent geometric series with common ratio $\left|\frac{x_1}{x_0}\right| < 1$.

Therefore, by comparison test, the series $\Sigma\left|a_n x_1^n\right|$ converges.

Hence, $\Sigma a_n x^n$ is absolutely convergent for every $x = x_1$, when $|x_1| < |x_0|$.

Note: The theorem asserts that if a power series converges for $x = x_0$ then it converges absolutely for all those x for which $|x| < |x_0|$. May be at $x = x_0$, the series converges but not absolutely.

It will be seen later that such a situation arises occasionally, when though absolutely convergent within the interval of convergence, the power series is just convergent at one or both the end points.

Theorem 3. *If a power series $\Sigma a_n x^n$ diverges for $x = x'$, then it diverges for every $x = x''$, where $|x''| > |x'|$.*

If the series was convergent for $x = x''$, then by Theorem 2 it would have to converge for all x with $|x| < |x''|$, and in particular at x', which contradicts the hypothesis. Hence the theorem.

In view of these two theorems, the radius of convergence is the supremum (least upper bound) of all the number $|x|$ for which $\Sigma\left|a_n x^n\right|$ converges. So we state that '*a power series is absolutely convergent within its interval of convergence and divergent outside it*'.

ILLUSTRATIONS

1. Radius of convergence of the series $\Sigma n x^{n-1}$ is $\dfrac{1}{\lim(n)^{1/n}} = 1$. So, the series converges absolutely on the interval $]-1, 1[$. The series does not converge at $x = \pm 1$.
2. The series $1 + x^2 + x^4 + x^6 + \ldots$, has unit radius of convergence.

CHAPTER 13

3. For the series $\sum_{n=1}^{\infty} \frac{x^{n-1}}{n}$, radius of convergence $R = \frac{1}{\overline{\lim}\left(\frac{1}{n}\right)^{1/n}} = 1$. It converges at $x = -1$ but not at $x = 1$. It converges absolutely on]–1, 1[.

4. The series $x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \dots$, has radius of convergence $R = 1/\overline{\lim}\left|\frac{1}{n}\right|^{1/n} = 1$.

   Hence, it converges absolutely on]–1, 1[. It converges at $x = 1$ but not at $x = -1$.

5. $\Sigma n^2 x^n$ has unit radius of convergence and therefore converges absolutely for $|x| < 1$. Does not converge at $x = \pm 1$.

Note: It is known that

$$\underline{\lim}_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right| \le \underline{\lim}_{n\to\infty}|a_n|^{1/n} \le \overline{\lim}_{n\to\infty}|a_n|^{1/n} \le \overline{\lim}_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right|$$

provided the first and the last limit exists.

Thus if $\underline{\lim}\left|\frac{a_{n+1}}{a_n}\right| = \overline{\lim}\left|\frac{a_{n+1}}{a_n}\right|$, then $\underline{\lim}|a_n|^{1/n} = \overline{\lim}|a_n|^{1/n}$, *i.e.*, if $\lim\left|\frac{a_{n+1}}{a_n}\right|$ exists, then $\lim|a_n|^{1/n}$ also exists and the two are equal.

So the radius of convergence can also be found by the relation

$$R = \lim_{n\to\infty}\left|\frac{a_n}{a_{n+1}}\right|$$

provided the limit exists, but recourse has to be made to $R = 1/\overline{\lim}|a_n|^{1/n}$ in case the former limit does not exist.

Example 1. Find the radius of convergence of the series

(*i*) $x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots$

(*ii*) $1 + x + 2!x^2 + 3!x^3 + 4!x^4 + \dots$

■ (*i*) The radius of convergence $R = \lim\left|\frac{a_n}{a_{n+1}}\right| = \lim\frac{(n+1)!}{n!} = \infty$. Therefore, the series converges absolutely for all values of x.

(*ii*) The radius of convergence $R = \lim_{n\to\infty}\frac{n!}{(n+1)!} = 0$. Therefore, the series converges for no value of x, of course other than zero.

Example 2. Find the radius of convergence of the series

$$\frac{1}{2}x+\frac{1\cdot 3}{2\cdot 5}x^2+\frac{1\cdot 3\cdot 5}{2\cdot 5\cdot 8}x^3+\ldots$$

■ Radius of convergence $R=\lim\limits_{n\to\infty}\dfrac{1\cdot 3\cdot 5\ldots(2n-1)\,2\cdot 5\cdot 8\ldots(3n+2)}{2\cdot 5\cdot 8\ldots(3n-1)\,1\cdot 3\cdot 5\ldots(2n+1)}=3/2.$

So, the series converges absolutely for all x, where $|x|<\dfrac{3}{2}$.

Example 3. Find the radius of convergence of the series

$$x+\frac{x^2}{2^2}+\frac{2!}{3^3}x^3+\frac{3!}{4^4}x^4+\ldots$$

■ Radius of convergence $R=\lim\dfrac{(n-1)!}{n^n}\dfrac{(n+1)^{n+1}}{n!}$

$$=\lim\left(1+\frac{1}{n}\right)^{n+1}=e$$

So, the series converges absolutely if $|x|<e$.

Example 4. Find the interval of absolute convergence for the series $\sum\limits_{n=1}^{\infty}x^n/n^n$.

■ It is a power series and will therefore be absolutely convergent within its interval of convergence. Now, the radius of convergence

$$R=\frac{1}{\overline{\lim}\,|a_n|^{1/n}}=\frac{1}{\lim\left|\dfrac{1}{n^n}\right|^{1/n}}=\infty$$

Hence, the series converges absolutely for all x.

3. PROPERTIES OF FUNCTIONS EXPRESSIBLE AS POWER SERIES

In this section, we shall derive some properties of the functions which can be expressed in terms of power series, *i.e.*, the functions of the form

$$f(x)=\sum_{n=0}^{\infty}a_nx^n,\text{ or } f(x)=\sum_{n=0}^{\infty}a_n(x-a)^n,$$

the former being the power series expansion of $f(x)$ about the origin, while the latter is about $x=a$. This can, however, be thought of in the reverse direction also. In the interval of convergence, the power series Σa_nx^n or $\Sigma a_n(x-a)^n$ has a definite sum $f(x)$ for each x, and usually different sum for a different x. In order to express this dependence on x, we write

$$f(x) = \Sigma a_n x^n \text{, or } f(x) = \Sigma a_n (x-a)^n,$$

$f(x)$ is then called the *sum function* of the series.

Before proceeding to the next theorem, let us understand an important distinction between the intervals of absolute and of uniform convergence. An interval of uniform convergence *must* include its end points but the interval of absolute convergence need not.

Thus, if a power series converges absolutely and uniformly for $|x| < R$, we express this fact by saying that it converges absolutely in $]-R, R[$, and uniformly in $[-R+\varepsilon, R-\varepsilon]$, no matter which $\varepsilon > 0$ is chosen; the latter interval may be replaced by $[-R_1, R_1]$, $R_1 < R$.

Theorem 4. *If a power series $\Sigma a_n x^n$ converges for $|x| < R$, and let us define a function*

$$f(x) = \Sigma a_n x^n, |x| < R,$$

then $\Sigma a_n x^n$ converges uniformly on $[-R+\varepsilon, R-\varepsilon]$, no matter which $\varepsilon > 0$ is chosen, and that the function f is continuous and differentiable on $]-R, R[$, and

$$f'(x) = \Sigma n a_n x^{n-1}, |x| < R \qquad ...(1)$$

Let $\varepsilon > 0$ be any number given.

For $|x| \le R - \varepsilon$, we have

$$|a_n x^n| \le |a_n| (R-\varepsilon)^n.$$

But since $\Sigma a_n (R-\varepsilon)^n$, converges absolutely (every power series converges absolutely within its interval of convergence), therefore by Weierstrass's M-test, the series $\Sigma a_n x^n$ converges uniformly on $[-R+\varepsilon, R-\varepsilon]$.

Again, since very term of the series $\Sigma a_n x^n$ is continuous and differentiable on $]-R, R[$, and $\Sigma a_n x^n$ is uniformly convergent on $[-R+\varepsilon, R-\varepsilon]$, therefore its sum function f is also continuous and differentiable on $]-R, R[$.

Also, $$\overline{\lim_{n\to\infty}} |n a_n|^{1/n} = \overline{\lim_{n\to\infty}} \left(n^{1/n}\right) |a_n|^{1/n} = 1/R$$

Hence, the differentiated series $\Sigma n a_n x^{n-1}$ is also a power series and has the same radius of convergence R as $\Sigma a_n x^n$. Therefore, $\Sigma n a_n x^{n-1}$ is uniformly convergent in $[-R+\varepsilon, R-\varepsilon]$.

Hence, $f'(x) = \Sigma n a_n x^{n-1}, |x| < R$.

Note: Since each term of the power series is continuous and integrable on $]-R, R[$ and the power series $\Sigma a_n x^n$ is uniformly convergent on $[-R+\varepsilon, R-\varepsilon]$, therefore the sum function f is continuous and integrable on the interval.

Moreover $\overline{\lim} \left|\dfrac{a_n}{n+1}\right|^{1/n} = 1/R$, so that the integrated series also has the same radius of convergence as $\Sigma a_n x^n$. Therefore the above theorem can include the case that f is integrable and the integrated series which also is a power series has the same radius of convergence. Moreover by repeated application of the theorem, f can be differentiated or integrated any number of times.

Corollary. *Under the hypothesis of the above theorem, f has derivatives of all orders in]–R, R[, which are given by*

$$f^{(m)}(x) = \sum_{n=m}^{\infty} n(n-1)\dots(n-m+1)a_n x^{n-m} \qquad \dots(2)$$

and in particular,

$$f^{(m)}(0) = m!a_m,\ m = 0, 1, 2, \dots \qquad \dots(3)$$

[Here, as usual, $f^{(m)}$ denotes the mth derivative of f for $m = 1, 2, 3, \dots$, and $f^{(0)}$ means f.]

Let $f(x) = a_0 + a_1 x + a_2 x^2 + \dots + a_n x^n + \dots$

By the above theorem, $f(x)$ is differentiable any number of times. Let us differentiate m times.

$$\therefore \quad f^{(1)}(x) = a_1 + 2a_2 x + 3a_3 x^2 + \dots + na_n x^{n-1} + \dots$$

$$f^{(2)}(x) = 2\cdot 1a_2 + 3\cdot 2a_3 x + 4\cdot 3a_4 x^2 + \dots + n(n-1)a_n x^{n-2} + \dots$$

$$f^{(3)}(x) = 3!a_3 + 4\cdot 3\cdot 2a_4 x + \dots + n(n-1)(n-2)a_n x^{n-3} + \dots$$

$$f^{(4)}(x) = 4!a_4 + \dots + n(n-1)(n-2)(n-3)a_n x^{n-4} + \dots$$

$$\vdots$$

$$f^{(m)}(x) = m!\,a_m + (m+1)m(m-1)\dots 2\cdot a_{m+1}\,x + \dots + n(n-1)\dots(n-m+1)a_n x^{n-m} + \dots$$

$$= \sum_{n=m}^{\infty} n(n-1)\dots(n-m+1)a_n x^{n-m}$$

Also

$$f^{(m)}(0) = m!a_m$$

the other terms vanishing at $x = 0$.

Remarks:

1. The formula (3) above is very interesting. On the one hand it shows that the coefficient of the power series expansion of f can be determined by the values at the origin of f and that of its derivatives. On the other hand, if the coefficients are given (*i.e.*, the power series is given), the values of f and its derivatives at the origin can be read off immediately from the power series.
2. A power series $\Sigma a_n x^n = f(x)$ convergent on an interval]–R, R[, $R \neq 0$ is in fact a Taylor's series for its sum function f, that is the one whose coefficients are determined by the Taylor's formula

$$a_m = \frac{f^{(m)}(0)}{m!},\ m = 0, 1, 2, \dots$$

Hence the coefficients of power series are uniquely specified by its sum. Thus, if a function f can be expanded into a power series convergent to the function, this series is Taylor's series for the function.

Now, it appears natural to pose the question whether the converse assertion is true. The problem can be stated as follows.

Suppose a function $f(x)$ is infinitely differentiable on an interval $]-R, R[$, $R \neq 0$. We can formally construct the Taylor's series for this function:

$$f(0)+\frac{f'(0)}{1!}x+\frac{f''(0)}{2!}x^2+\ldots+\frac{f^{(n)}(0)}{n!}x^n+\ldots$$

Now does this series converge on the interval $]-R, R[$, and will its sum be equal to the function f in case it exists? It turns out that in general the answer to the question is negative which can be confirmed by the example of the function

$$f(x)=\begin{cases} e^{-1/x^2}, & \text{for } x \neq 0 \\ 0, & \text{for } x=0 \end{cases}$$

In fact it can be easily verified (Ex. 5, Page 187) that the function is infinitely differentiable through the x-axis and that at the origin, we have

$$f(0)=f'(0)=\ldots=f^{(n)}(0)=\ldots=0$$

Consequently, all the coefficients of the Taylor's series of the function are equal to zero. Thus the Taylor's series converges on the entire x-axis and its sum is identically equal to zero, whereas the function takes on a zero value only at the origin and so we fail to express f as a power series.

4. ABEL'S THEOREM

In this section we shall prove that for a power series which has a given radius of convergence and convergent at an end-point of the interval, the interval of uniform convergence extends up to and includes that end-point. Moreover, in that case the sum function f is continuous not only in $]-R, R[$, but also at the end point. This is proved in Abel's Theorem:

Abel's Theorem (*First form*). *If a power series* $\sum_{n=0}^{\infty} a_n x^n$ *converges at the end point* $x = R$ *of the interval of convergence* $]-R, R[$, *then it is uniformly convergent in the closed interval* $[0, R]$.

We shall show that under the assumptions of the theorem, Cauchy's criterion for uniform convergence is satisfied on the closed interval $[0, R]$. This will imply the uniform convergence of the series on $[0, R]$.

Let $S_{n,p}=a_{n+1}R^{n+1}+a_{n+2}R^{n+2}+\ldots+a_{n+p}R^{n+p}$, $p=1, 2, \ldots$

Then obviously

$$\begin{aligned} a_{n+1}R^{n+1} &= S_{n,1} \\ a_{n+2}R^{n+2} &= S_{n,2}-S_{n,1} \\ &\vdots \\ a_{n+p}R^{n+p} &= S_{n,p}-S_{n,p-1} \end{aligned} \qquad \ldots(1)$$

Let $\varepsilon > 0$ be given.

Since the number series $\sum_{n=0}^{\infty} a_n R^n$ is convergent, therefore by Cauchy's General Principle of convergence, there exists an integer N such that for $n \geq N$,

$$\left| S_{n,q} \right| < \varepsilon, \text{ for all } q = 1, 2, 3, \ldots \qquad \ldots(2)$$

Taking into account that

$$\left(\frac{x}{R}\right)^{n+p} \leq \left(\frac{x}{R}\right)^{n+p-1} \leq \ldots \leq \left(\frac{x}{R}\right)^{n+1} \leq 1, \text{ for } 0 \leq x \leq R$$

and using equations (1) and (2), we have for $n \geq N$

$$\begin{aligned}
&\left| a_{n+1}x^{n+1} + a_{n+2}x^{n+2} + \ldots + a_{n+p}x^{n+p} \right| \\
&= \left| a_{n+1}R^{n+1}\left(\frac{x}{R}\right)^{n+1} + a_{n+2}R^{n+2}\left(\frac{x}{R}\right)^{n+2} + \ldots a_{n+p}R^{n+p}\left(\frac{x}{R}\right)^{n+p} \right| \\
&= \left| S_{n,1}\left\{\left(\frac{x}{R}\right)^{n+1} - \left(\frac{x}{R}\right)^{n+2}\right\} + S_{n,2}\left\{\left(\frac{x}{R}\right)^{n+2} - \left(\frac{x}{R}\right)^{n+3}\right\} + \ldots \right. \\
&\quad \left. + S_{n,p-1}\left\{\left(\frac{x}{R}\right)^{n+p-1} - \left(\frac{x}{R}\right)^{n+p}\right\} + S_{n,p}\left(\frac{x}{R}\right)^{n+p} \right| \\
&\leq \left| S_{n,1} \right| \left\{\left(\frac{x}{R}\right)^{n+1} - \left(\frac{x}{R}\right)^{n+2}\right\} + \left| S_{n,2} \right| \left\{\left(\frac{x}{R}\right)^{n+2} - \left(\frac{x}{R}\right)^{n+3}\right\} + \ldots \\
&\quad + \left| S_{n,p-1} \right| \left\{\left(\frac{x}{R}\right)^{n+p-1} - \left(\frac{x}{R}\right)^{n+p}\right\} + \left| S_{n,p} \right| \left(\frac{x}{R}\right)^{n+p} \\
&< \varepsilon\left\{\left(\frac{x}{R}\right)^{n+1} - \left(\frac{x}{R}\right)^{n+2} + \left(\frac{x}{R}\right)^{n+2} - \left(\frac{x}{R}\right)^{n+3} + \ldots - \left(\frac{x}{R}\right)^{n+p} + \left(\frac{x}{R}\right)^{n+p}\right\} \\
&= \varepsilon\left(\frac{x}{R}\right)^{n+1} \leq \varepsilon.
\end{aligned}$$

for all $n \geq N$, $p \geq 1$ and for all $x \in [0, R]$.

Hence by Cauchy's criterion, the series converges uniformly on $[0, R]$.

Remark: Since a series of continuous functions that uniformly converges in a given interval, defines a continuous function in that interval, it follows that a power series (with radius of convergence R) which converges at the end-point R, defines a *continuous function* in the interval $[-R + \varepsilon, R]$.

Notes:

1. We have a similar situation in case a power series with interval of convergence $]-R, R[$, converges at the left end $x=-R$, or at both the end points; in the former case the series is uniformly convergent on $[-R, 0]$ and in the latter on $[-R, R]$.
2. If a power series with interval of convergence $]-R, R[$ diverges at the end point $x = R$, it cannot be uniformly convergent on the interval $[0, R]$. For otherwise, if the series is uniformly convergent on $[0, R]$, it will converge at $x = R$ as well, which contradicts the given condition.

Abel's Theorem (*Second form*). *If $\Sigma a_x x^n$ be a power series with finite radius of convergence R, and let*

$$f(x) = \Sigma a_n x^n, \; -R < x < R$$

If the series $\Sigma a_n R^n$ converges, then

$$\lim_{x \to R-0} f(x) = \Sigma a_n R^n$$

Let us first show that there is no loss of generality in taking $R = 1$.

Put $x = Ry$, so that

$$\Sigma a_n x^n = \Sigma a_n R^n y^n = \Sigma b_n y^n, \text{ where } b_n = a_n R^n.$$

It is a power series with radius of convergence R', where

$$R' = \frac{1}{\overline{\lim} \left| a_n R^n \right|^{1/n}} = 1$$

Thus, it will suffice to prove the following:

Let $\sum_0^{\infty} a_n x^n$ be a power series with unit radius of convergence and let

$$f(x) = \sum_0^{\infty} a_n x^n, \; -1 < x < 1$$

If the series Σa_n converges, then

$$\lim_{x \to 1-0} f(x) = \sum_0^{\infty} a_n$$

Let $S_n = a_0 + a_1 + a_2 + \ldots + a_n$, $S_{-1} = 0$, and let $\sum_{n=0}^{\infty} a_n = S$, then

$$\sum_{n=0}^{m} a_n x^n = \sum_{n=0}^{m} \left(S_n - S_{n-1}\right) x^n = \sum_{n=0}^{m-1} S_n x^n + S_m x^m - \sum_{n=0}^{m} S_{n-1} x^n$$

$$= \sum_{n=0}^{m-1} S_n x^n - x \sum_{n=0}^{m} S_{n-1} x^{n-1} + S_m x^m = (1-x) \sum_{n=0}^{m-1} S_n x^n + S_m x^m .$$

For $|x| < 1$, when $m \to \infty$, since $S_m \to S$, and $x^m \to 0$, we get

$$f(x) = (1-x)\sum_{n=0}^{\infty} S_n x^n, \text{ for } 0 < x < 1 \qquad \text{...(1)}$$

Again, since $S_n \to S$, for $\varepsilon > 0$, there exists N such that

$$|S_n - S| < \varepsilon/2, \text{ for all } n \geq N \qquad \text{...(2)}$$

Also

$$(1-x)\sum_{n=0}^{\infty} x^n = 1, |x| < 1 \qquad \text{...(3)}$$

Hence, for $n \geq N$, we have, for $0 < x < 1$,

$$|f(x) - S| = \left|(1-x)\sum_{n=0}^{\infty} S_n x^n - S\right| \qquad \text{[using (1)]}$$

$$= \left|(1-x)\sum_{n=0}^{\infty}(S_n - S)x^n\right| \qquad \text{[using (3)]}$$

$$\leq (1-x)\sum_{n=0}^{N}|S_n - S|\,x^n + \frac{\varepsilon}{2}(1-x)\sum_{n=N+1}^{\infty} x^n \qquad \text{[using (2)]}$$

$$\leq (1-x)\sum_{n=0}^{N}|S_n - S|\,x^n + \frac{\varepsilon}{2}$$

But for a fixed N, $(1-x)\sum_{n=0}^{N}|S_n - S|\,x^n$ is a positive continuous function of x, having zero value at $x = 1$. Therefore, there exists $\delta > 0$, such that for $1 - \delta < x < 1$, $(1-x)\sum_{n=0}^{N}|S_n - S|\,x^n < \varepsilon/2$.

$$\therefore\ |f(x) - S| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon, \text{ when } 1 - \delta < x < 1$$

Hence, $\lim_{x \to 1-0} f(x) = S = \sum_{n=0}^{\infty} a_n$.

Corollary. *If the series* $\Sigma(-1)^n a_n$ *converges, then*

$$\lim_{x \to -1+0} f(x) = \Sigma(-1)^n a_n.$$

Putting $y = -x$, and $b_n = (-1)^n a_n$, we have

$$\lim_{x \to -1+0} f(x) = \lim_{x \to -1+0} \Sigma a_n x^n = \lim_{x \to -1+0} \Sigma(-1)^n a_n (-x)^n$$

$$= \lim_{y \to 1-0} \Sigma b_n y^n = \Sigma b_n.$$

Remarks:

1. The power series $\Sigma a_n x^n$ has radius of convergence R, and $\Sigma a_n R^n$ is also convergent, therefore by Abel's test for uniform convergence, the series $\Sigma a_n x^n$ converges uniformly in $[-R+\varepsilon, R]$ *i.e.*, the interval of uniform convergence extends upto and includes the end-point R.
2. The theorem implies that the sum function f is continuous at the end-point R.
   Since the power series converges uniformly in $[-R+\varepsilon, R-\varepsilon]$, and each term of the series is continuous, therefore the sum function f is also continuous on $[-R+\varepsilon, R-\varepsilon]$. As by the above theorem f is also continuous at $x = R$, therefore f is continuous (in fact uniformly) on $[-R+\varepsilon, R-\varepsilon]$.

An application. *If Σa_n, Σb_n, Σc_n converge to the sums A, B, C, respectively, and if*

$$c_n = a_0 b_n + a_1 b_{n-1} + \ldots + a_n b_0, \textit{ then}$$

$$\Sigma a_n \cdot \Sigma b_n = \Sigma c_n \text{ or } AB = C$$

Let $$f(x) = \sum_{n=0}^{\infty} a_n x^n,\ g(x) = \sum_{n=0}^{\infty} b_n x^n,\ h(x) = \sum_{n=0}^{\infty} c_n x^n, \text{ for } 0 \le x \le 1.$$

For $|x| < 1$, the three series converge absolutely, therefore $\Sigma c_n x^n$ is the *Cauchy product* of $\Sigma a_n x^n$ and $\Sigma b_n x^n$, and so

$$f(x) \cdot g(x) = h(x),\ 0 \le x < 1 \qquad \text{...(4)}$$

Also by Abel's theorem,

$$\left.\begin{aligned} &\lim_{x \to 1-0} f(x) = \sum_{0}^{\infty} a_n \Rightarrow f(x) \to A \text{ as } x \to 1-0 \\ \text{Similarly, } &g(x) \to B,\ h(x) \to C \qquad \text{as } x \to 1-0 \end{aligned}\right\} \qquad \text{...(5)}$$

Hence, (4) and (5) imply, $AB = C$.

4.1 Taylor's Theorem

Let $\sum_{n=0}^{\infty} a_n x^n$ *be a power series with radius of convergence R, and let*

$$f(x) = \sum_{n=0}^{\infty} a_n x^n,\ |x| < R$$

Then for any $a \in]-R, R[$, prove that f can be expanded in a power series about 'a' which converges for $|x - a| < R - |a|$, and

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n,\ |x-a| < R - |a|$$

Suppose $|x-a|<R-|a|$. Then $|x|\le|x-a|+|a|<R$, and thus $\Sigma a_n x^n$ converges.

Now,
$$f(x)=\sum_{n=0}^{\infty}a_n x^n=\sum_{n=0}^{\infty}a_n\left(\overline{x-a}+a\right)^n$$
$$=\sum_{n=0}^{\infty}a_n\sum_{m=0}^{n}\binom{n}{m}a^{n-m}(x-a)^m \quad \text{...(1)}$$

We wish to change the order of summation in this expression. To prove its validity we notice that it is the summation by rows of a double series, which if convergent absolutely, will converge by columns as well to the same sum.

Replacing all quantities by their moduli and taking all terms with positive sign, expression on the right of (1) gives
$$\sum_{n=0}^{\infty}|a_n|\sum_{m=0}^{n}\binom{n}{m}|a_n|^{n-m}|x-a|^m=\sum_{n=0}^{\infty}|a_n|\cdot\left(|x-a|+|a|\right)^n$$
which is a power series, and converges, since $|x-a|+|a|<R$

Hence, for $|x-a|<R-|a|$, change in the order of summation is justified and so
$$f(x)=\sum_{m=0}^{\infty}\left\{\sum_{n=m}^{\infty}\binom{n}{m}a_n a^{n-m}\right\}(x-a)^m \quad \text{...(2)}$$

This is the required result, but for the coefficients which we shall now express in terms of the values of f and its derivatives at a point.

We know
$$f(x)=a_0+a_1x+a_2x^2+\ldots+a_mx^m+\ldots+a_nx^n+\ldots$$
$$\therefore\quad f^{(1)}(x)=a_1+2a_2x+3a_3x^2+\ldots+ma_mx^{m-1}+\ldots+na_nx^{n-1}+\ldots$$
$$f^{(2)}(x)=2!a_2+3\cdot 2a_3x+\ldots+m(m-1)a_mx^{m-2}+\ldots+n(n-1)a_nx^{n-2}+\ldots$$
$$\vdots$$
$$f^{(m)}(x)=m!\,a_m+(m+1)m(m-1)\ldots3\cdot 2a_{m+1}x+\ldots+n(n-1)\ldots(n-m+1)a_nx^{n-m}+\ldots$$
$$=m!\left[a_m+\binom{m+1}{m}a_{m+1}x+\binom{m+2}{m}a_{m+2}x^2+\ldots\right]=m!\sum_{n=m}^{\infty}\binom{n}{m}a_nx^{n-m}$$
$$\therefore\quad f^{(m)}(a)=m!\sum_{n=m}^{\infty}\binom{n}{m}a_na^{n-m}.$$

Hence from (2),
$$f(x)=\sum_{m=0}^{\infty}\frac{f^{(m)}(a)}{m!}(x-a)^m,\ \text{for } |x-a|<R-|a|.$$

CHAPTER 13

Example 5. Show that

(*i*) $\tan^{-1} x = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \ldots, -1 \le x \le 1$

(*ii*) $\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \ldots,$ (Gregory's series)

(*iii*) $\frac{1}{2}\left(\tan^{-1} x\right)^2 = \frac{x^2}{2} - \frac{x^4}{4}\left(1 + \frac{1}{3}\right) + \frac{x^6}{6}\left(1 + \frac{1}{3} + \frac{1}{5}\right) + \ldots, -1 < x \le 1.$

■ (*i*) We know

$$\left(1 + x^2\right)^{-1} = 1 - x^2 + x^4 - x^6 + \ldots, |x| < 1$$

The series on the right is a power series with radius of convergence 1; so that it is convergent, absolutely in]–1, 1[, and uniformly in $[-k, k], |k| < 1$. The integrated series is also convergent absolutely in]–1, 1[and uniformly in $[-k, k], |k| < 1$.

Integrating, we get

$$\tan^{-1} x = C + x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \ldots, |x| < 1$$

where C is constant of integration, which can be seen to be equal to zero by putting $x = 0$.

$$\therefore \qquad \tan^{-1} x = x - \frac{x^3}{5} + \frac{x^5}{5} - \ldots, |x| < 1$$

Series on the right is a power series with radius of convergence equal to 1.

However, the series $x - \frac{x^3}{3} + \frac{x^5}{5} - \ldots$ is convergent at $x = \pm 1$ also, so that by Abel's theorem, it is uniformly convergent in [–1, 1], and hence

$$\tan^{-1} x = x - \frac{x^3}{3} + \frac{x^5}{5} - \ldots, -1 \le x \le 1$$

At $x = 1$, we get by Abel's theorem (second form),

(*ii*) $\frac{\pi}{4} = \tan^{-1} 1 = \lim_{x \to 1-0} \tan^{-1} x = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \ldots,$

(*iii*) We have

$$\tan^{-1} x = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \ldots, -1 \le x \le 1$$

and

$$\left(1 + x^2\right)^{-1} = 1 - x^2 + x^4 - x^6 + \ldots, -1 < x < 1.$$

Both the series are absolutely convergent in $]-1, 1[$, therefore their Cauchy product will converge absolutely to the product of their sums, $(1+x^2)^{-1}\tan^{-1}x$ in $]-1, 1[$.

$$\therefore \qquad (1+x^2)^{-1}\tan^{-1}x = x - \left(1+\frac{1}{3}\right)x^3 + \left(1+\frac{1}{3}+\frac{1}{5}\right)x^5 - \ldots, \quad -1<x<1$$

Integrating,

$$\frac{1}{2}\left(\tan^{-1}x\right)^2 = \frac{x^2}{2} - \frac{x^4}{4}\left(1+\frac{1}{3}\right) + \frac{x^6}{6}\left(1+\frac{1}{3}+\frac{1}{5}\right) - \ldots, \quad -1<x<1$$

the constant of integration vanishes.

It can be easily seen that the power series of the right converges at $x = 1$ also, so that by Abel's theorem,

$$\frac{1}{2}\left(\tan^{-1}x\right)^2 = \frac{x^2}{2} - \frac{x^4}{4}\left(1+\frac{1}{3}\right) + \frac{x^6}{6}\left(1+\frac{1}{3}+\frac{1}{5}\right) - \ldots, -1<x\le 1$$

Example 6. Show that

$$\log(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \ldots, -1<x\le 1$$

and deduce that

$$\log 2 = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \ldots$$

■ We know

$$(1+x)^{-1} = 1 - x + x^2 - x^3 + \ldots, -1<x<1$$

Integrating,

$$\log(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \ldots, -1<x<1$$

the constant of integration vanishes as can be verified by putting $x = 0$.

The power series on the right converges at $x = 1$ also. Therefore by Abel's theorem

$$\log(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \ldots, -1<x\le 1$$

At $x = 1$, we get, by Abel's theorem (second form),

$$\log 2 = \lim_{x\to 1-}\log(1+x) = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \ldots$$

Example 7. Show that

$$\frac{1}{2}\left[\log(1+x)\right]^2 = \frac{x^2}{2} - \frac{x^3}{3}\left(1+\frac{1}{2}\right) + \frac{x^4}{4}\left(1+1\frac{1}{2}+\frac{1}{3}\right) - \ldots, -1<x\le 1$$

- We know

$$\log(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \dots, -1 < x \le 1$$

and
$$(1+x)^{-1} = 1 - x + x^2 - x^3 + x^4 - \dots, -1 < x < 1$$

Both the series are absolutely convergent in]–1, 1[, therefore their Cauchy product will converge to $(1 + x)^{-1} \log (1 + x)$. Thus

$$(1+x)^{-1}\log(1+x) = x - x^2\left(1+\frac{1}{2}\right) + x^3\left(1+\frac{1}{2}+\frac{1}{3}\right) - \dots, -1 < x < 1$$

Integrating,

$$\frac{1}{2}\left[\log(1+x)\right]^2 = \frac{x^2}{2} - \frac{x^3}{3}\left(1+\frac{1}{2}\right) + \frac{x^4}{4}\left(1+\frac{1}{2}+\frac{1}{3}\right) - \dots, -1 < x < 1$$

the constant of integration vanishes.

Since the series on the right converges at $x = 1$, also, therefore by Abel's Theorem, we have

$$\frac{1}{2}\left[\log(1+x)\right]^2 = \frac{x^2}{2} - \frac{x^3}{3}\left(1+\frac{1}{2}\right) + \frac{x^4}{4}\left(1+\frac{1}{2}+\frac{1}{3}\right) - \dots, -1 < x \le 1$$

EXERCISE

1. Determine the radius of convergence, and the exact interval of convergence of each of the following power series:

   (*i*) $\Sigma \frac{(n+1)x^n}{(n+2)(n+3)}$, (*ii*) $\Sigma \frac{nx^n}{(n+1)^2}$, (*iii*) $\Sigma \frac{2^n x^n}{n!}$,

   (*iv*) $\Sigma \frac{(n!)^2 x^{2n}}{(2n)!}$, (*v*) $\Sigma \frac{(-1)^n x^{2n}}{(n!)^2 2^{2n}}$, (*vi*) $\Sigma 3^{-n} x^{3n}$,

   (*vii*) $\Sigma \frac{(x-1)^n}{2^n}$, (*viii*) $\Sigma \frac{(-1)^{n+1}}{n}(x-1)^n$, (*ix*) $\Sigma \frac{n!\,(x+2)^n}{n^n}$,

   (*x*) $\sum_{n=2}^{\infty} \frac{(x+2)^n}{\log n}$, (*xi*) $\Sigma \frac{1\cdot 2\cdot 3 \dots n}{1\cdot 3\cdot 5 \dots (2n-1)} x^{2n}$, (*xii*) $\Sigma (-1)^n \frac{x^{2n+1}}{(2n+1)!}$.

2. Prove that the power series

$$1 + \frac{a\cdot b}{1\cdot c}x + \frac{a(a+1)b(b+1)}{1\cdot 2c(c+1)}x^2 + \dots,$$

   has unit radius of convergence.

3. If the power series $\Sigma a_n x^n$ has radius of convergence R, then prove that, for any positive integer k, $\Sigma a_n x^{kn}$ has radius of convergence $R^{1/k}$.

4. Discuss the uniform convergence with respect to x of the series, $\sum_{n=1}^{\infty} \frac{x^n}{n^{\alpha}}$.

5. Show that the following power series are uniformly convergent:

(*i*) $1+\frac{x}{2}+\frac{x^2}{3}+\frac{x^4}{4}+..., [-1, K], 0<K<1,$

(*ii*) $x+\frac{x^2}{2^2}+\frac{x^3}{3^2}+\frac{x^4}{4^2}+..., [-1, 1],$

(*iii*) $x+\frac{1}{2}\frac{x^3}{3}+\frac{1\cdot 3}{2\cdot 4}\frac{x^5}{5}+\frac{1\cdot 3\cdot 5}{2\cdot 4\cdot 6}\frac{x^7}{7}+..., [-1, 1].$

6. Show that

(*i*) $\log(1-x)=-x-\frac{x^2}{2}-\frac{x^3}{3}-..., -1\le x<1,$

(*ii*) $\log 2=1-\frac{1}{2}+\frac{1}{3}-\frac{1}{4}+...,$

(*iii*) $\frac{1}{2}[\log(1-x)]^2=\frac{x^2}{2}+\left(1+\frac{1}{2}\right)\frac{x^3}{3}+\left(1+\frac{1}{2}+\frac{1}{3}\right)\frac{x^4}{4}+..., -1\le x<1.$

7. Show that

(*i*) $\sin^{-1}x=x+\frac{1}{2}\frac{x^3}{3}+\frac{1\cdot 3}{2\cdot 4}\frac{x^5}{5}+\frac{1\cdot 3\cdot 5}{2\cdot 4\cdot 6}\frac{x^7}{7}+..., -1<x\le 1,$

(*ii*) $\frac{\pi}{2}=1+\frac{1}{2}\cdot\frac{1}{3}+\frac{1\cdot 3}{2\cdot 4}\cdot\frac{1}{5}+\frac{1\cdot 3\cdot 5}{2\cdot 4\cdot 6}\cdot\frac{1}{7}+...,$

(*iii*) $\frac{1}{2}\left(\sin^{-1}x\right)^2=\frac{x^2}{2}+\frac{2}{3}\cdot\frac{x^4}{4}+\frac{2\cdot 4}{3\cdot 5}\cdot\frac{x^6}{6}+..., -1<x\le 1.$

Also obtain similar expression for

$\cos^{-1}x, \sinh^{-1}x, \cosh^{-1}x, \tanh^{-1}x$

8. Prove that

$$\int_0^x\frac{dt}{1+t^n}=x-\frac{x^{n+1}}{n+1}+\frac{x^{2n+1}}{2n+1}-..., -1<x\le 1, n>0.$$

ANSWERS

1. (*i*) $R=1, [-1, 1[$ (*ii*) $R=1, [-1, 1[$

(*iii*) $R=\infty, \mathbf{R}$ (*iv*) $R=2,]-2, 2[$

(*v*) $R=\infty, \mathbf{R}$ (*vi*) $R=3^{1/3},]-R, R[$

(*vii*) $R=2,]-1, 3[$ (*viii*) $R=1,]0, 2]$

(*ix*) $R=e,]-2-e, -2+e[$ (*x*) $R=1, [-3, -1[$

(*xi*) $R=\sqrt{2},]-R, R[$ (*xii*) $R=1, [-1, 1].$

CHAPTER 13

14 Fourier Series

1. TRIGONOMETRICAL SERIES

One of the most important and also one of the most interesting fields to which we may apply the subject developed so far, is provided by the theory of *Fourier series* and more generally by that of *trigonometrical series*, which we now propose to study. Such series were first obtained in theoretical physics, in the course of investigations on periodic motion, chiefly in acoustics, optics, electrodynamics and the theory of heat. The credit to initiate a thorough study of certain trigonometrical series goes to Fourier, although he did not go far enough to discover the fundamental results of the theory.

Periodic processes (motions) are described mathematically by means of periodic functions. A function f of one independent variable x is said to be periodic if there exists $\lambda \neq 0$ (called its period) such that $f(x+\lambda) = f(x)$, for all values of x. The trigonometric functions, $\sin x$, $\cos x$ are the simplest periodic functions with period 2π.

A series of the form

$$\frac{1}{2}a_0 + (a_1 \cos x + b_1 \sin x) + (a_2 \cos 2x + b_2 \sin 2x) + \ldots$$

or more compactly

$$\frac{1}{2}a_0 + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \sin nx)$$

is called a *trigonometric series*, and the constants a_0, a_n, b_n ($n = 1, 2, 3 \ldots$) are called its coefficients. If the series converges for all values of x (in fact, in an interval $c \leq x \leq c + 2\pi$), then in view of the periodicity of the trigonometric functions, its sum is a function f defined for all values of x and periodic with period 2π. To save tedious repetition, it may be mentioned once for all that the *functions considered in this chapter are all assumed to be periodic.*

We now proceed to obtain a relationship between the function f and the coefficients a_n, b_n which was conjectured by *Euler*.

The following results of integral calculus form the basis of all subsequent work. Throughout, we use *m, n to denote positive integers or zero.*

$$\int_{-\pi}^{\pi} \cos mx \cos nx \, dx = \begin{cases} 0, & \text{for } m \neq n \\ \pi, & \text{for } m = n > 0 \\ 2\pi, & \text{for } m = n = 0 \end{cases}$$

$$\int_{-\pi}^{\pi} \sin mx \sin nx \, dx = \begin{cases} 0, \text{ for } m \neq n \text{ and } m = n = 0 \\ \pi, \text{ for } m = n > 0 \end{cases}$$

$$\int_{-\pi}^{\pi} \cos mx \sin nx \, dx = 0$$

1.1 Fourier Series

If the numbers $a_0, a_1, \ldots, a_n, \ldots, b_1, b_2, \ldots, b_n, \ldots$ are derived from a function f by means of *Euler-Fourier formulas*:

$$\begin{aligned} a_n &= \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos nx \, dx, \; n = 0, 1, 2, \ldots \\ b_n &= \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin nx \, dx, \; n = 0, 1, 2, \ldots \end{aligned} \qquad \ldots(1)$$

then the series

$$\frac{1}{2}a_0 = \sum_{n=1}^{\infty}(a_n \cos nx + b_n \sin nx) \qquad \ldots(2)$$

is called the *Fourier series of* f or the *Fourier series generated by* f, and the coefficients a_n, b_n defined by (1) as the *Fourier coefficients** of f.

It is to be noted that the Fourier coefficients have been obtained purely on the assumption that the function f is bounded and integrable on $[-\pi, \pi]$.

There is nothing to suggest that the Fourier series (2) is convergent. In fact the series may not converge at all, and even if it converges, the sum may not be f, though it often is and there is some justification for the hope that the series may converge and have f for its sum. In case the Fourier series of f converges *uniformly*, the definitions of Fourier constants suggest that its sum will be f, and that f is capable of a unique Fourier series expansion.

A host of questions arise but there are four major problems which prick the mind and which we shall try to answer in the subsequent pages.

1. Is the Fourier series of a given (integrable) function f convergent for some or all values of x in $-\pi \le x \le \pi$?
2. If it converges, does the Fourier series of f converge to f?
3. When and how can a function be expressed as a Fourier series?
4. When the trigonometric series be the Fourier series?

CHAPTER 14

* Justification for the definitions can be seen from the following:

Let the series $\frac{1}{2}a_0 + \Sigma(a_n \cos nx + b_n \sin nx)$ converge uniformly on $[-\pi, \pi]$ to the sum function f.

Multiplying by $\cos nx$ (since uniformity of convergence of the series is not thus destroyed, term-by-term integration is justified) and integrating, we get

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f \cos nx \, dx$$

Other terms are vanishing in view of the results of integral calculus mentioned earlier.

Similarly, $$a_0 = \frac{1}{\pi}\int_{-\pi}^{\pi} f \, dx, \; b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f \sin nx \, dx$$

Note: One of the answer to the last question is straightforward, since we know that, if the trigonometric series

$$\frac{1}{2}a_0 + \sum_{n=1}^{\infty}(a_n \cos nx + b_n \sin nx)$$

is uniformly convergent on $[-\pi, \pi]$ and if $f(x)$ is its sum, then it is the Fourier series of $f(x)$.

Further, since

$$|a_n \cos nx + b_n \sin nx| \leq |a_n \cos nx| + |b_n \sin nx| \leq |a_n| + |b_n|.$$

Therefore, by Weierstrass's M-test, it follows that, *if the series* $\sum_{n=1}^{\infty}(|a_n| + |b_n|)$ *converges, then the trigonometric series converges absolutely and uniformly on every closed interval* $[a, b]$ *and that it is the Fourier series of a continuous* 2π*-periodic function.*

2. SOME PRELIMINARY THEOREMS

Before dealing directly with the problems raised above, let us do some preliminary theorems which will help in the subsequent development of the subject.

2.1 Periodic Functions

We know that a periodic function of period λ is such that $f(x \pm \lambda) = f(x)$, for all x.

Theorem 1. *For a periodic function of period* 2π, *prove that*

(*i*) $\int_{\alpha}^{\beta} f\,dx = \int_{\alpha+2\pi}^{\beta+2\pi} f\,dx,$

(*ii*) $\int_{-\pi}^{\pi} f\,dx = \int_{\alpha}^{\alpha+2\pi} f\,dx,$

(*iii*) $\int_{-\pi}^{\pi} f(x)\,dx = \int_{-\pi}^{\pi} f(\gamma + x)\,dx,$

α, β, γ *being any numbers whatsoever.*

(*i*) For a periodic function of period 2π, we know

$$f(t - 2\pi) = f(t)$$

Hence, putting $t = x + 2\pi$, for all α and β, we get

$$\int_{\alpha}^{\beta} f(x)dx = \int_{\alpha+2\pi}^{\beta+2\pi} f(t - 2\pi)dt = \int_{\alpha+2\pi}^{\beta+2\pi} f(t)dt = \int_{\alpha+2\pi}^{\beta+2\pi} f(x)dx$$

(*ii*) $$\int_{\alpha}^{\alpha+2\pi} f\,dx = \int_{\alpha}^{-\pi} f\,dx + \int_{-\pi}^{\pi} f\,dx + \int_{\pi}^{\alpha+2\pi} f\,dx$$

$$= \int_{\alpha}^{-\pi} f\,dx + \int_{-\pi}^{\pi} f\,dx + \int_{-\pi}^{\alpha} f\,dx \qquad \text{[using } (i)\text{]}$$

$$= \int_{-\pi}^{\pi} f\,dx$$

(*iii*) Let $\gamma + x = t$

$$\therefore \qquad \int_{-\pi}^{\pi} f(\gamma + x)dx = \int_{\gamma-\pi}^{\gamma+\pi} f(t)dt$$

$$= \int_{\gamma-\pi}^{-\pi} f\,dt + \int_{-\pi}^{\pi} f\,dt + \int_{\pi}^{\gamma+\pi} f\,dt$$

$$= \int_{\gamma+\pi}^{-\pi} f\,dt + \int_{-\pi}^{\pi} f\,dt + \int_{\pi}^{\gamma+\pi} f\,dt \qquad \text{[using } (i)\text{]}$$

$$= \int_{-\pi}^{\pi} f\,dt = \int_{-\pi}^{\pi} f\,dx.$$

These results, in fact, mean that the *integral of a periodic function over any interval whose length is equal to its period always has the same value*.

2.2 Some Definitions

Piecewise monotonic functions. A function f is called *piecewise monotonic* on an interval $[a, b]$, if there exists a partition of $[a, b]$ such that the function is monotonic (that is either monotone increasing or monotone decreasing) on each of the sub-intervals. *Piecewise continuous functions* are defined in the same way.

From the definition, it follows that if the function f is piecewise monotonic and bounded on the interval $[a, b]$, then it can have discontinuities of the first kind (or *finite discontinuous*) only. Indeed, if $x = c$ is a point of discontinuity of the function, then by virtue of the monotonicity of f, there exist the limits

$$\lim_{x \to c-} f(x) = f(c-), \ \lim_{x \to c+} f(x) = f(c+)$$

i.e., the point c is a *finite discontinuity* or a discontinuity of the first kind.

A point c is a point of *infinite discontinuity* of a function f if

$$\lim_{x \to c} f(x) = \pm\infty$$

Dirichlet's Integral. Integrals of the following two forms are called Dirichlet's integrals

$$\int_0^a f \frac{\sin nx}{\sin x} dx, \ \int_0^a f \frac{\sin nx}{x} dx.$$

2.3 Some Theorems

Theorem 2. *If f is bounded and integrable on $[-\pi, \pi]$ and if a_n, b_n are its Fourier coefficients, then $\sum_{n=1}^{\infty} \left(a_n^2 + b_n^2\right)$ converges.*

The integral

$$\int_{-\pi}^{\pi}\left(f - \sum_{n=1}^{m}\left(a_n \cos nx + b_n \sin nx\right)\right)^2 dx \geq 0$$

as its integrand is never negative, now

$$\int_{-\pi}^{\pi}\left[f - \sum_{n=1}^{m}\left(a_n \cos nx + b_n \sin nx\right)\right]^2 dx$$

$$= \int_{-\pi}^{\pi} f^2 dx + \int_{-\pi}^{\pi}\left[\sum_{n=1}^{m}\left(a_n \cos nx + b_n \sin nx\right)\right]^2 dx$$

$$- 2\Sigma\left[a_n \int_{-\pi}^{\pi} f \cos nx\, dx\right] - 2\Sigma\left[b_n \int_{-\pi}^{\pi} f \sin nx\, dx\right]$$

$$= \int_{-\pi}^{\pi} f^2 dx + \pi\,\Sigma a_n^2 + \pi\,\Sigma b_n^2 - 2\pi\,\Sigma a_n^2 - 2\pi\,\Sigma b_n^2$$

$$= \int_{-\pi}^{\pi} f^2 dx - \pi \sum_{n=1}^{m}\left(a_n^2 + b_n^2\right)$$

Since this is non-negative, we have

$$\sum_{n=1}^{m}\left(a_n^2 + b_n^2\right) \leq \frac{1}{\pi}\int_{-\pi}^{\pi} f^2\, dx,$$

taking limit as $m \to \infty$, we obtain

$$\sum_{n=1}^{\infty}\left(a_n^2 + b_n^2\right) \leq \frac{1}{\pi}\int_{-\pi}^{\pi} f^2\, dx \quad \textbf{(Bessel's inequality)}$$

Hence, the series $\sum_{n=1}^{\infty}\left(a_n^2 + b_n^2\right)$ is convergent, since f, and so f^2, is integrable on $[-\pi, \pi]$.

Corollary. Since $\Sigma\left(a_n^2 + b_n^2\right)$ converges, its general term $\left(a_n^2 + b_n^2\right) \to 0$, and so, $a_n \to 0$, $b_n \to 0$. Thus the *Fourier coefficients a_n and b_n of an integrable (and bounded) function form a null sequence, i.e., when $n \to \infty$,*

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f \cos nx\, dx \to 0, \text{ and } b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f \sin nx\, dx \to 0.$$

Note: If the function is *piecewise continuous* on the interval $[-\pi, \pi]$ then also its Fourier coefficients approach zero as $n \to \infty$.

For, then a partition of $[-\pi, \pi]$ exists in each sub-interval of which f and so f^2 is continuous and integrable, and therefore the integral over $[-\pi, \pi]$ can be expressed as a sum of definite integrals of continuous functions over the sub-intervals.

Theorem 3. ***Riemann-Lebesgue.*** *If a function ϕ is bounded and integrable on the interval $[a, b]$, then as $n \to \infty$,*

$$A_n = \int_a^b \phi \cos nx \, dx \to 0, \text{ and } B_n = \int_a^b \phi \sin nx \, dx \to 0.$$

If a and b both belong to one and the same interval of the form $-m\pi \le x \le m\pi$ (m, any positive integer), we define $f(x) = \phi(x)$ in $a \le x \le b$, and $f(x) = 0$ at the remaining points of $[-m\pi, m\pi]$; for other real x, f is defined so as to be periodic with period 2π. Then

$$A_n = \int_a^b \phi \cos nx \, dx = \int_{-\pi}^{\pi} f \cos nx \, dx = \pi a_n,$$

and, similarly $B_n = \pi b_n$, where a_n and b_n denote the Fourier coefficients of f.

Using Theorem 2, A_n, B_n tend to zero.

Again, if a and b do not lie in the same interval of the type $[-m\pi, m\pi]$, we can split up the interval $[a, b]$ into a finite number of sub-intervals, each of which lies in an interval of the form $[-m\pi, m\pi]$. A_n and B_n can be expressed then as the sum of a (fixed) finite number of terms each of which tends to 0, as $n \to \infty$. Thus A_n and B_n tend to 0.

Deduction. For a bounded integrable function ϕ, limits of the Dirichlet's integrals are equal, *i.e.*,

$$\lim \int_0^a \phi \frac{\sin nx}{\sin x} dx = \lim \int_0^a \phi \frac{\sin nx}{x} dx, \; 0 \le a < \pi$$

Assigning the value 0 to $\left(\frac{1}{\sin x} - \frac{1}{x}\right)$ at the origin, the function becomes continuous and bounded in $[0, a]$, $a < \pi$. Thus the function $\phi\left(\frac{1}{\sin x} - \frac{1}{x}\right)$ is bounded and integrable on $[0, a]$.

Therefore by the above theorem

$$\lim_{n\to\infty} \int_0^a \phi\left(\frac{1}{\sin x} - \frac{1}{x}\right) \sin nx \, dx = 0$$

or

$$\lim \int_0^a \phi \frac{\sin nx}{\sin x} dx = \lim \int_0^a \phi \frac{\sin nx}{x} dx$$

Note: Compare with Example 16.

Theorem 4. *If a function f is bounded and integrable in* $[0, a]$, $a > 0$, *and monotone in* $]0, \delta[$, $0 < \delta < a$, *then*

$$\lim_{n\to\infty} \int_0^\infty f \frac{\sin nx}{x} dx = f(0+) \int_0^\infty \frac{\sin x}{x} dx$$

Let us first consider the case when $f(0+) = 0$. We may suppose, as it does not affect the result that $f(0) = 0$.

CHAPTER 14

Let δ be a positive number less than a.

By the second mean value theorem there exists $\delta_1 \in [0, \delta]$ such that

$$\int_0^{\delta} f\frac{\sin nx}{x}dx = f(0)\int_0^{\delta_1}\frac{\sin nx}{x}dx + f(\delta)\int_{\delta_1}^{\delta}\frac{\sin nx}{x}dx$$

$$= f(\delta)\int_{\delta_1}^{\delta}\frac{\sin nx}{x}dx$$

$$= f(\delta)\int_{n\delta_1}^{n\delta}\frac{\sin t}{t}dt = f(\delta)\int_{n\delta_1}^{n\delta}\frac{\sin x}{x}dx \qquad ...(1)$$

Since $\int_0^{\infty}\frac{\sin x}{x}dx$ is convergent, there exists $k > 0$ such that for all $x \geq 0$,

$$\left|\int_0^{X}\frac{\sin x}{x}dx\right| \leq k$$

$$\Rightarrow \qquad \left|\int_{n\delta_1}^{n\delta}\frac{\sin x}{x}dx\right| = \left|\int_0^{n\delta}\frac{\sin x}{x}dx - \int_0^{n\delta_1}\frac{\sin x}{x}dx\right| \leq 2k \qquad ...(2)$$

Also $f(\delta) \to 0$, as $\delta \to 0+0$, therefore for any $\varepsilon > 0$, there exists $\gamma > 0$ such that

$$|f(\delta)| < \varepsilon/4k, \text{ when } 0 < \delta < \gamma \qquad ...(3)$$

Hence from equations (1), (2) and (3), we get

$$\left|\int_0^{\delta} f\frac{\sin nx}{x}dx\right| \leq |f(\delta)|\cdot 2k < \frac{1}{2}\varepsilon, \text{ when } 0 < \delta < \gamma$$

Now, writing

$$\int_0^{a} f\frac{\sin nx}{x}dx = \int_0^{\delta} f\frac{\sin nx}{x}dx + \int_{\delta}^{a} f\frac{\sin nx}{x}dx$$

and observing that (by Theorem 3) the second integral on the right $\to 0$, as $n \to \infty$, we see that there exists a positive integer m such that for all $n \geq m$,

$$\left|\int_0^{a} f\frac{\sin nx}{x}dx\right| < \frac{1}{2}\varepsilon + \frac{1}{2}\varepsilon = \varepsilon$$

$$\Rightarrow \qquad \lim_{n\to\infty}\int_0^{a} f\frac{\sin nx}{x}dx = 0$$

so that the theorem is proved for the case when $f(0+) = 0$.

In the *general case*, when $f(0+) \neq 0$, since

$$\lim_{x\to 0+}\left[f(x) - f(0+)\right] = 0$$

replacing f by $[f - f(0+)]$, we get

$$\Rightarrow \qquad \lim_{n\to\infty}\int_0^a \left|f - f(0+)\right|\frac{\sin nx}{x}dx = 0$$

$$\lim_{n\to\infty}\int_0^a f\frac{\sin nx}{x}dx = \lim_{n\to\infty} f(0+)\int_0^a \frac{\sin nx}{x}dx$$

$$= f(0+)\lim_{n\to\infty}\int_0^{na}\frac{\sin t}{t}dt, \ (t = nx)$$

$$= f(0+)\int_0^{\infty}\frac{\sin x}{x}dx$$

Hence, the theorem.

Deductions

1. Since $\int_0^{\infty}\frac{\sin x}{x}dx = \frac{\pi}{2}$

$$\therefore \qquad \lim_{n\to\infty}\int_0^a f\frac{\sin nx}{x}dx = f(0+)\int_0^{\infty}\frac{\sin x}{x}dx = \frac{\pi}{2}f(0+)$$

2. Using deduction Theorem 3, we get

$$\lim_{n\to\infty}\int_0^a f\frac{\sin nx}{x}dx = f(0+)\int_0^{\infty}\frac{\sin x}{x}dx = \frac{\pi}{2}f(0+)$$

Theorem 5. *If f is bounded and integrable in $[-\pi, \pi]$ and monotonic in $[-\delta, 0[$ and $]0, \delta]$, (not necessarily in the same sense), where $0 < \delta < \pi$, then*

$$\frac{1}{2}a_0 = \sum_{n=1}^{\infty} a_n = \frac{f(0-)+f(0+)}{\pi}\int_0^{\infty}\frac{\sin x}{x}dx$$

where a_n, $n = 0, 1, 2, \ldots$, *denote the Fourier's coefficients of f.*

Now,

$$\frac{1}{2}a_0 + \sum_{n=1}^{m} a_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} f\,dx + \frac{1}{\pi}\sum_{n=1}^{m}\int_{-\pi}^{\pi} f\cos nx\,dx$$

$$= \frac{1}{\pi}\int_{-\pi}^{\pi} f\left[\frac{1}{2} + \sum_{n=1}^{m}\cos nx\right]dx$$

$$= \frac{1}{\pi}\int_{-\pi}^{\pi} f\frac{\sin\left(m+\frac{1}{2}\right)x}{2\sin\frac{1}{2}x}dx$$

$$= \frac{1}{2\pi}\int_0^{\pi} f(-x)\frac{\sin\left(m+\frac{1}{2}\right)x}{\sin\frac{1}{2}x}dx + \frac{1}{2\pi}\int_0^{\pi} f\frac{\sin\left(m+\frac{1}{2}\right)x}{\sin\frac{1}{2}x}dx$$

$$= \frac{1}{2\pi} 2\int_0^{\pi/2} f(-2x)\, \frac{\sin(2m+1)x}{\sin x} dx + \frac{1}{2\pi} 2\int_0^{\pi/2} f(2x) \frac{\sin(2m+1)x}{\sin x} dx$$

Hence letting $m \to \infty$,

$$\frac{1}{2}a_0 + \sum_{n=1}^{\infty} a_n = \frac{1}{\pi}\left[f(0-) + f(0+)\right] \int_0^{\infty} \frac{\sin x}{x} dx$$

Corollary. Using $\int_0^{\infty} \frac{\sin x}{x} dx = \frac{\pi}{2}$, we get

$$\frac{1}{2}a_0 + \sum_{n=1}^{\infty} a_n = \frac{f(0-) + f(0+)}{2}$$

Deduction. The theorem may be used to deduce the value of the integral $\int_0^{\infty} \frac{\sin x}{x} dx$.

Taking $f(x) = 1,\ \forall\, x$, we get

$$\frac{1}{2}a_0 = \frac{1}{2\pi} \int_{-\pi}^{\pi} 1 \cdot dx = 1$$

and
$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} \cos nx\, dx = 0$$

so that

$$1 = \frac{f(0-) + f(0+)}{\pi} \int_0^{\infty} \frac{\sin x}{x} dx = \frac{2}{\pi} \int_0^{\infty} \frac{\sin x}{x} dx$$

$\Rightarrow$
$$\int_0^{\infty} \frac{\sin x}{x} dx = \frac{1}{2}\pi$$

3. THE MAIN THEOREM

We are now in a position to take up the main problem. Our approach will be to write down the partial sum sequence of the Fourier series of the function f and lay down a set of *sufficient conditions* so that the sequence (and consequently the Fourier series) converges to f. That will answer all the first three questions raised earlier in § 1.1, that a function f satisfying these conditions will generate a Fourier series which converges and has a sum f.

A large number of sets of sufficient conditions for the purpose have been found, of which we shall use only one, **Dirichlet's criterion**, the great generality of which renders it sufficient for most purposes. It was given by Dirichlet and was the first set of exact conditions of convergence in the theory of Fourier series.

Theorem 6. *If a function f is bounded periodic with period 2π and integrable on $[-\pi, \pi]$, and piecewise monotonic on $[-\pi, \pi]$, then*

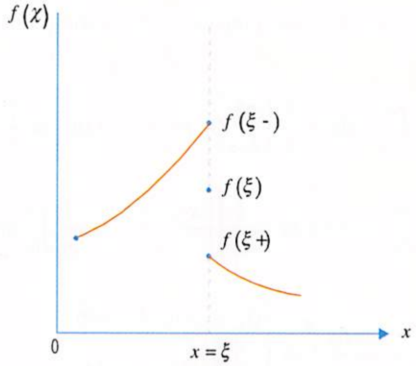
$$\frac{1}{2}a_0 + \sum_{n=1}^{\infty}\left(a_n \cos n\xi + b_n \sin n\xi\right) = \begin{cases} \frac{1}{2}\left[f(\xi-) + f(\xi+)\right], \text{ for } -\pi < \xi < \pi, \\ \frac{1}{2}\left[f(\pi-) + f(-\pi+)\right], \text{ for } \xi = \pm\pi \end{cases}$$

where a_n, b_n are Fourier coefficients of f.

Let $\frac{1}{2}a_0 + \sum_{n=1}^{\infty}\left(a_n \cos nx + b_n \sin nx\right)$ be the Fourier series of f, and ξ, a point of $[-\pi, \pi]$.

The mth partial sum at the point ξ,

$$\begin{aligned}
&\frac{1}{2}a_0 + \sum_{n=1}^{m}\left(a_n \cos n\xi + b_n \sin n\xi\right) \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} f\,dx + \sum_{n=1}^{m}\frac{1}{\pi}\int_{-\pi}^{\pi} f\left[\cos nx\ \cos n\xi + \sin nx \sin n\xi\right]dx \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} f\left[1 + 2\sum_{n=1}^{m}\cos n(x-\xi)\right]dx \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x+\xi)\left[1 + 2\sum_{n=1}^{m}\cos nx\right]dx \qquad (\S 2.1) \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x+\xi)\frac{\sin\left(m+\frac{1}{2}\right)x}{\sin\frac{1}{2}x}dx \\
&= \frac{1}{2\pi}\int_{-\pi}^{0} f(x+\xi)\frac{\sin\left(m+\frac{1}{2}\right)x}{\sin\frac{1}{2}x}dx + \frac{1}{2\pi}\int_{0}^{\pi} f(x+\xi)\frac{\sin\left(m+\frac{1}{2}\right)x}{\sin\frac{1}{2}x}dx \\
&= \frac{1}{\pi}\int_{0}^{\pi/2} f(-2t+\xi)\frac{\sin(2m+1)t}{\sin t}dt + \frac{1}{\pi}\int_{0}^{\pi/2} f(2t+\xi)\frac{\sin(2m+1)t}{\sin t}dt
\end{aligned}$$



Discontinuity at ξ

Fig. 1

Proceeding to limits when $m \to \infty$,

$$\frac{1}{2}a_0 + \sum_{n=1}^{\infty}(a_n \cos n\xi + b_n \sin n\xi) = \frac{1}{\pi}\left[\frac{1}{2}\pi f(\xi-) + \frac{1}{2}\pi f(\xi+)\right]$$

$$= \frac{f(\xi-) + f(\xi+)}{2}, \quad \xi \in [-\pi, \pi]$$

Thus, the Fourier series of a (periodic) function f which is bounded, integrable and piecewise monotonic on $[-\pi, \pi]$, converges to $\frac{1}{2}[f(\xi-) + f(\xi+)]$ at a point $\xi, -\pi < \xi < \pi$, and (using periodicity of f) to $\frac{1}{2}[f(\pi-) + f(-\pi+)]$ at the ends, $\pm\pi$.

Corollary 1. At a point ζ of continuity of f, the sum of the Fourier series of f,

$$\frac{1}{2}a_0 + \sum_{n=1}^{\infty}(a_n \cos n\zeta + b_n \sin n\zeta) = \frac{1}{2}[f(\zeta-) + f(\zeta+)]$$

$$= \frac{1}{2}[f(\zeta) + f(\zeta)] = f(\zeta)$$

Thus for a function f satisfying the conditions of the theorem, *the sum of the Fourier series is actually $f(x)$ at all points x where f is continuous*; while at points of (finite) discontinuity the sum of the series is $\frac{1}{2}(f(x-) + f(x+)]$.

Corollary 2. Since a function of bounded variation can be expressed as a difference of two monotone increasing function (Jordan theorem), it may be easily shown that a *periodic function of bounded variation can be expressed as a Fourier series.*

Example 1. Find the Fourier series of the periodic function f with period 2π, defined as follows:

$$f(x) = \begin{cases} 0, \text{ for } -\pi < x \le 0, \\ x, \text{ for } 0 \le x \le \pi \end{cases}$$

What is the sum of the series at $x = 0, \pm\pi, 4\pi, -5\pi$?

■ The function is bounded, integrable and piecewise monotonic on $[-\pi, \pi]$. Let us determine the Fourier coefficients

$$a_0 = \frac{1}{\pi}\int_{-\pi}^{\pi} f\,dx = \frac{1}{\pi}\left[\int_{-\pi}^{0} 0 \cdot dx + \int_{0}^{\pi} x\,dx\right] = \frac{\pi}{2}$$

$$a_n = \frac{1}{\pi}\int_{0}^{\pi} x\cos nx\,dx = \frac{1}{\pi}\left[\left|\frac{x\sin nx}{n}\right|_0^{\pi} - \frac{1}{n}\int_0^{\pi}\sin nx\,dx\right]$$

$$= \frac{1}{\pi n}\left|\frac{\cos nx}{n}\right|_0^{\pi} = \begin{cases} 0, & \text{for } n \text{ even} \\ -\dfrac{2}{\pi n^2}, & \text{for } n \text{ odd} \end{cases}$$

$$b_n = \frac{1}{\pi}\int_0^{\pi} x \sin nx\, dx = -\frac{\cos n\pi}{n} = \begin{cases} -\dfrac{1}{n}, & \text{for } n \text{ even} \\[2ex] \dfrac{1}{n}, & \text{for } n \text{ odd} \end{cases}$$

In $[-\pi, \pi]$, the points $\pm\pi$ are the only points of discontinuity of *f*. Therefore at all points of the interval other than $\pm\pi$, we have

$$f(x) = \frac{\pi}{4} - \frac{2}{\pi}\left\{\frac{\cos x}{1^2} + \frac{\cos 3x}{3^2} + \frac{\cos 5x}{5^2} + \ldots\right\} + \left\{\frac{\sin x}{1} - \frac{\sin 2x}{2} + \frac{\sin 3x}{3} - \ldots\right\}$$

At $x = 0$ where *f* is continuous, we get

$$f(0) = \frac{\pi}{4} - \frac{2}{\pi}\left\{\frac{1}{1^2} + \frac{1}{3^2} + \frac{1}{5^2} + \ldots\right\}$$

$$\Rightarrow \qquad 1 + \frac{1}{3^2} + \frac{1}{5^2} + \ldots = \frac{\pi^2}{8}$$

Because of periodicity, value of the Fourier series at $x = 4\pi$ is same as at $x = 0$.

At $x = \pm\pi$, points of discontinuity of *f*, the sum of the series

$$= \frac{1}{2}\left[f(\pi -) + f(-\pi +)\right] = \frac{1}{2}\pi$$

$$\therefore \quad \frac{\pi}{2} = \frac{\pi}{4} + \frac{2}{\pi}\left\{\frac{1}{1^2} + \frac{1}{3^2} + \frac{1}{5^2} + \ldots\right\}$$

or

$$1 + \frac{1}{3^2} + \frac{1}{5^2} + \ldots = \frac{\pi^2}{8}$$

Again because of periodicity, value at -5π is same as at $x = \pm\pi$.

Note: Figure 2 gives the shape of the function *f*.

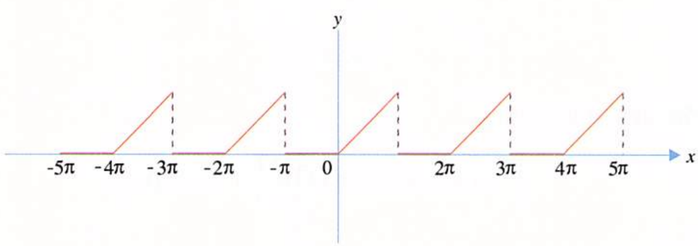


Fig. 2

Example 2. Expand in a series of sines and cosines of multiple angles of *x*, the periodic function *f* with period 2π is defined as

$$f(x) = \begin{cases} -1, & \text{for } -\pi < x < 0 \\ 1, & \text{for } 0 \le x \le \pi \end{cases}$$

Also calculate the sum of the series at $x = 0, \frac{1}{2}\pi, \pm\pi$.

■ The function (Fig. 3) is piecewise monotonic, bounded and integrable on $[-\pi, \pi]$. Let us compute its Fourier coefficients

$$a_0 = \frac{1}{\pi}\int_{-\pi}^{\pi} f\,dx = \frac{1}{\pi}\left[\int_{-\pi}^{0}(-1)\,dx + \int_{0}^{\pi} 1 \cdot dx\right] = 0$$

$$a_n = \frac{1}{\pi}\left[\int_{-\pi}^{0} -\cos nx\,dx + \int_{0}^{\pi} \cos nx\,dx\right] = 0$$

$$b_n = \frac{1}{\pi}\left[\int_{-\pi}^{0} -\sin nx\,dx + \int_{0}^{\pi} \sin nx\,dx\right] = \frac{2}{\pi n}(1-\cos n\pi) = \begin{cases} 0, & \text{for } n \text{ even} \\ \dfrac{4}{\pi n}, & \text{for } n \text{ odd} \end{cases}$$

The function is continuous at all points of $[-\pi, \pi]$ except $\pm\pi$.

$$\therefore\ f(x) = \frac{4}{\pi}\left\{\frac{\sin x}{1} + \frac{\sin 3x}{3} + \frac{\sin 5x}{5} + \ldots\right\}$$

which holds at all points with the exception of the discontinuities, $\pm\pi$.

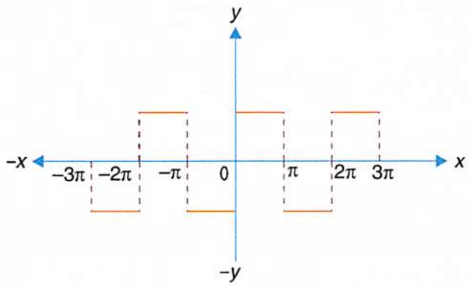


Fig. 3

At $x = \pm\pi$, the sum of the series

$$= \frac{1}{2}\left[f(\pi-) + f(-\pi+)\right] = \frac{1}{2}(1-1) = 0$$

At $x = \frac{1}{2}\pi$, a point of continuity,

$$1 = \frac{4}{\pi}\left\{1 - \frac{1}{3} + \frac{1}{5} - \ldots\right\}$$

$$\therefore \qquad 1 - \frac{1}{3} + \frac{1}{5} - \ldots = \frac{\pi}{4}$$

Note: Figure 3 is the graph of the given function.

Figure 4 illustrates how the partial sums S_m of the series represent more and more accurately the function f as $m \to \infty$.

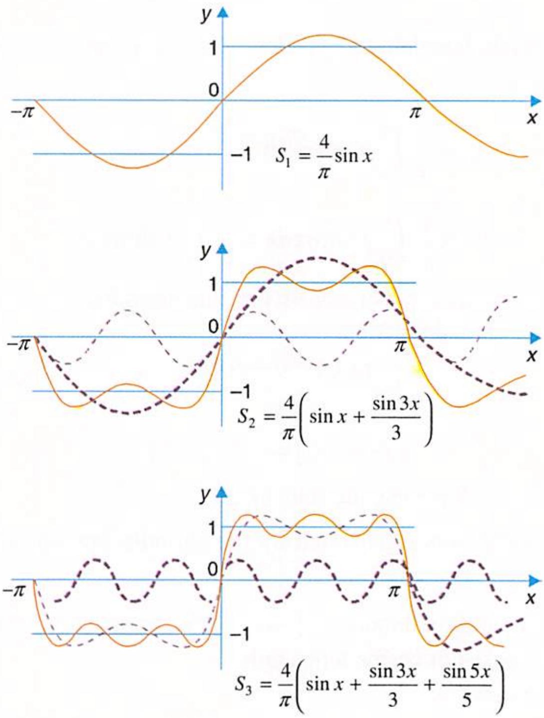


Fig. 4

3.1 Fourier Series for Even and Odd Functions

Even Functions. If f is an even function, *i.e.*, $f(-x) = f(x)$, $\forall\ x$, then $f\cos nx$ is an even and $f\sin nx$ is an odd function and therefore

$$a_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f\cos nx\,dx = \frac{1}{\pi}\left[\int_{-\pi}^{0} f\cos nx\,dx + \int_{0}^{\pi} f\cos nx\,dx\right] = \frac{2}{\pi}\int_{0}^{\pi} f\cos nx\,dx \qquad \text{...(1)}$$

$$b_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f\sin nx\,dx = 0.$$

So, the Fourier series of an even function consists of terms of cosines only and the coefficients a_n may be computed from (1).

Also, since for an even function,

$$f(0+) = f(0-) = f(0)$$

and

$$f(-\pi+0)=f(\pi-0)$$

the sum of the series is $f(0)$ at 0 or $\pm$(even multiple of π), and is $f(\pi-)$ at $\pm\pi$ or $\pm$(odd multiple of π).

Odd Functions. If f is an odd function, *i.e.*, $f(-x)=-f(x)$, $\forall\ x$, then $f\cos nx$ is an odd and $f\sin nx$ an even function and therefore

$$a_n=\frac{1}{\pi}\int_{-\pi}^{\pi} f\cos nx\,dx=0$$

$$b_n=\frac{1}{\pi}\int_{-\pi}^{\pi} f\sin nx\,dx=\frac{2}{\pi}\int_{0}^{\pi} f\sin nx\,dx \qquad ...(2)$$

So the Fourier series of an odd function consists of sine terms only and the coefficients b_n may be calculated from (2). Also for an odd function

$$f(0-)=0=f(0+)$$

and

$$f(-\pi+0)=-f(\pi-0)$$

The sum of the series is 0 at $x=0$ (or any multiple of π).

Example 3. Find the Fourier series generated by the periodic function $|x|$ of period 2π. Also compute the value of series at $0,\ 2\pi,\ -3\pi$.

■ The function is monotone and continuous on $[-\pi,\pi]$. Moreover it is an even function and therefore the Fourier series will consist of cosine terms only.

The function may be restated as

$$f(x)=\begin{cases}-x, & \text{for } -\pi\le x\le 0\\ x, & \text{for } 0\le x\le\pi\end{cases}$$

$$a_0=\frac{1}{\pi}\left[\int_{-\pi}^{0}-x\,dx+\int_{0}^{\pi}x\,dx\right]=\frac{2}{\pi}\int_{0}^{\pi}x\,dx=\pi$$

$$a_n=\frac{2}{\pi}\int_{0}^{\pi}x\cos nx\,dx=\frac{2}{\pi n^2}(\cos n\pi-1)$$

$$=\begin{cases}0, & \text{for } n \text{ even}\\ -\dfrac{4}{\pi n^2}, & \text{for } n \text{ odd}\end{cases}$$

We, thus, obtain the series

$$f(x)=\frac{\pi}{2}-\frac{4}{\pi}\left[\frac{\cos x}{1^2}+\frac{\cos 3x}{3^2}+\frac{\cos 5x}{5^2}+...\right]$$

The series converges at all points and its sum is equal to the given function.

Sum of the series at 2π is the same as at 0.

At 0,

$$0 = \frac{\pi}{2} - \frac{4}{\pi}\left[\frac{1}{1^2} + \frac{1}{3^2} + \frac{1}{5^2} + \ldots\right]$$

$$\Rightarrow \qquad 1 + \frac{1}{3^2} + \frac{1}{5^2} + \ldots = \frac{\pi^2}{8}$$

At -3π (same as at π),

$$\pi = \frac{\pi}{2} + \frac{4}{\pi}\left[1 + \frac{1}{3^2} + \frac{1}{5^2} + \ldots\right]$$

$$\Rightarrow \qquad 1 + \frac{1}{3^2} + \frac{1}{5^2} + \ldots = \frac{\pi^2}{8}, \text{ same as above.}$$

3.2 Half Range Series

With the help of the Main theorem and those of even and odd functions, we now consider the expansion of a function over the interval $[0, \pi]$ in terms of (*i*) sine terms only, (*ii*) cosine terms only.

(*i*) ***The sine series.*** *If a function f is bounded, integrable and piecewise monotonic in* $[0, \pi]$, *then the sum of the sine series*

$$\Sigma b_n \sin nx, \text{ where } b_n = \frac{2}{\pi}\int_0^{\pi} f \sin nx\, dx$$

is equal to $\frac{1}{2}[f(x-) + f(x+)]$ *at every point x between* 0 *and* π, *and is equal to* 0, *when* $x = 0,\ \pi$.

To obtain a series consisting of only sine terms we define an odd function F in $[-\pi, \pi]$, identical with f in $[0, \pi]$.

Let $F = f$ in $[0, \pi]$, and $F(x) = -F(-x) = -f(-x)$ in $[-\pi, 0]$.

Evidently, F is bounded, integrable and piecewise monotone in $[-\pi, \pi]$, (*i.e.* satisfies the conditions of the main theorem). Again, F being an odd function, its Fourier series consists of sine terms only with sum

$$\Sigma b_n \sin nx, \text{ where } b_n = \frac{2}{\pi}\int_0^{\pi} F \sin nx\, dx$$

$$= \frac{2}{\pi}\int_0^{\pi} f \sin nx\, dx$$

equal to $\frac{1}{2}[F(x-) + F(x+)] = \frac{1}{2}[f(x-) + f(x+)]$ at every point x between 0 and π, and equal to 0 at $x = 0,\ \pi$.

(*ii*) ***The cosine series.*** *If f is bounded, integrable and piecewise monotonic in $[0, \pi]$, then the sum of the cosine series*

$$\frac{1}{2}a_0 + \Sigma a_n \cos nx, \text{ where } a_n = \frac{2}{\pi}\int_0^{\pi} | f \cos nx | \, dx$$

is equal to $\frac{1}{2}[f(x-) + f(x+)]$ *at every point x between 0 and π, and $f(0+)$ at $x = 0$, and $f(\pi-)$ at $x = \pi$.*

To prove the result, define an even function F on $[-\pi, \pi]$, identical with f on $[0, \pi]$, so that

$$F = f \text{ on } [0, \pi] \text{ and } F(x) = F(-x) = f(-x) \text{ on } [-\pi, 0].$$

Proceeding as above, we get the required result.

Example 4. Find the Fourier series consisting of (*i*) sine terms only, (*ii*) cosine terms only, which represents the periodic function $f(x) = x$ in $0 \le x \le \pi$.

■ (*i*) ***The sine series.*** The function may be extended as an odd function, $f(x) = x$ in $-\pi < x < \pi$, periodic with period 2π.

$\therefore$ $$a_n = 0, \text{ for } n = 0, 1, 2, 3, \ldots$$

and

$$b_n = \frac{2}{\pi}\int_0^{\pi} x \sin nx \, dx$$

$$= -\frac{2\cos n\pi}{n} = \begin{cases} -2/n, & \text{for } n \text{ even} \\ 2/n, & \text{for } n \text{ odd} \end{cases}$$

Hence for all points between 0 and π,

$$x = 2\left[\frac{\sin x}{1} - \frac{\sin 2x}{2} + \frac{\sin 3x}{3} - \ldots\right]$$

According to § 3.2, the sum of the series must be 0 at $x = 0$, π and this fact can be verified directly as well. The representation holds at $x = 0$, but not at $x = \pi$.

(*ii*) ***The cosine series.*** The function may be extended as an even periodic function $f(x) = |x|$ in $[-\pi, \pi]$ with period 2π.

$\therefore$ $$b_n = 0, \text{ for } n = 1, 2, 3, \ldots$$

and

$$a_0 = \frac{2}{\pi}\int_0^{\pi} x \, dx = \pi$$

$$a_n = \frac{2}{\pi}\int_0^{\pi} x \cos nx \, dx = \begin{cases} 0, & \text{for } n \text{ even} \\ -4/\pi n^2, & \text{for } n \text{ odd} \end{cases}$$

Hence for all points between 0 and π,

$$x = \frac{1}{2}\pi - \frac{4}{\pi}\left\{\frac{\cos x}{1^2} + \frac{\cos 3x}{3^2} + \frac{\cos 5x}{5^2} + \ldots\right\}$$

The function being continuous, the relation holds for all x.

According to § 3.2, the sum of the series must be $f(0+) = 0$ at $x = 0$, and $f(\pi -) = \pi$ at $x = \pi$ which can be found directly from the above relation.

At $x = 0$ or π, we get

$$1 + \frac{1}{3^2} + \frac{1}{5^2} + \ldots = \frac{\pi^2}{8}$$

4. INTERVALS OTHER THAN $[-\pi, \pi]$

So far we have considered the interval $[-\pi, \pi]$ only. It was just a matter of convenience, otherwise any finite interval could have been used. We now show that by effecting certain transformations, any finite interval can be made to correspond to the interval $[-\pi, \pi]$.

4.1 The Interval $[0, 2\pi]$

If f is bounded, integrable and piecewise monotonic in $[0, 2\pi]$, *then the sum of the series*

$$\frac{1}{2}a_0 + \sum_{n=1}^{\infty}(a_n \cos nx + b_n \sin nx),$$

where $a_n = \frac{1}{\pi}\int_0^{2\pi} f \cos nx\, dx$, $b_n = \frac{1}{\pi}\int_0^{2\pi} f \sin nx\, dx$ *is* $\frac{1}{2}[f(x-) + f(x+)]$ *at every point x between* 0 *and* 2π, *and is* $\frac{1}{2}[f(2\pi -) + f(0+)]$ *at* $x = 0, 2\pi$ *and is periodic with period* 2π.

On substituting $x = y + \pi$, we find that y varies from $-\pi$ to π as x varies from 0 to 2π, and

$$f(x) = f(y + \pi) = F(y), \text{ say}$$

Since f is bounded, integrable and piecewise monotonic in $[0, 2\pi]$, then so is F in $[-\pi, \pi]$. Hence by Dirichlet's criterion (Main theorem) the sum of the series $\frac{1}{2}a_0' + \sum_{n=1}^{\infty}(a_n' \cos ny + b_n' \sin ny)$, where $a_n' = \frac{1}{\pi}\int_{-\pi}^{\pi} F \cos ny\, dy$, $b_n' = \frac{1}{\pi}\int_{-\pi}^{\pi} F \sin ny\, dy$,

is $\frac{1}{2}[F(y-) + F(y+)]$ at every point y between $-\pi$ and π,

and is $\frac{1}{2}[F(\pi -) + F(-\pi +)]$ at $y = \pm\pi$, and is periodic with period 2π.

On changing the variable, we see that

$$a'_n = \frac{1}{\pi}\int_0^{2\pi} F(x-\pi)\cos n(x-\pi)\,dx = \frac{(-1)^n}{\pi}\int_0^{2\pi} f(x)\cos nx\,dx$$

$$b'_n = \frac{1}{\pi}\int_0^{2\pi} F(x-\pi)\ \sin n(x-\pi)\,dx = \frac{(-1)^n}{\pi}\int_0^{2\pi} f(x)\sin nx\,dx$$

Also

$$\frac{1}{2}[F(y-)+F(y+)] = \frac{1}{2}[f(x-)+f(x+)]$$

and

$$\frac{1}{2}[F(\pi-)+F(-\pi+)] = \frac{1}{2}[f(2\pi-)+f(0+)]$$

On making these changes, we get the required result.

4.2 Interval [–*l*, *l*], *l* is a Real Number

If f is bounded, integrable and piecewise monotonic in [–l, l], then the sum of the series

$$\frac{1}{2}a_0 + \Sigma\left[a_n\cos\frac{n\pi x}{l} + b_n\sin\frac{n\pi x}{l}\right],$$

where $a_n = \frac{1}{l}\int_{-l}^{l} f\cos\frac{n\pi x}{l}dx,\ b_n = \frac{1}{l}\int_{-1}^{1} f\sin\frac{n\pi x}{l}dx$, is $\frac{1}{2}[f(x-)+f(x+)]$ *for every x between –l and l, and is* $\frac{1}{2}[f(l-)+f(-l+)]$ *for* $x = \pm l$, *and is periodic with period 2l.*

On making the substitution $y = \pi x/l$, we see that y varies from $-\pi$ to π, as x varies from $-l$ to l, and

$$f(x) = f\left(\frac{yl}{\pi}\right) = F(y), \text{ say}$$

The function F satisfies the conditions of the Main Theorem in $[-\pi, \pi]$, and therefore proceeding as in the previous section (§ 4.1) we may prove the required result.

Example 5. The function x^2 is periodic with period $2l$ on the interval $[-l, l]$. Find its Fourier series.

■ The substitution $y = \pi x/l$ transforms the function into a periodic function with period 2π on $[-\pi, \pi]$. Moreover it is an even function.

$\therefore$ $$b_n = 0,\ n = 1, 2, 3, \ldots$$

$$a_0 = \frac{2}{l}\int_0^l x^2 dx = \frac{2l^2}{3}$$

$$a_n = \frac{2}{l}\int_0^l x^2\cos\frac{n\pi x}{l}dx = \frac{4l^2}{\pi^2 n^2}\cos n\pi$$

$$= \begin{cases} \dfrac{4l^2}{\pi^2 n^2}, & \text{for } n \text{ even} \\ -\dfrac{4l^2}{\pi^2 n^2}, & \text{for } n \text{ odd} \end{cases}$$

Also, the function is continuous on $[-l, l]$. Therefore for all x,

$$x^2 = \frac{l^2}{3} + \frac{4l^2}{\pi^2}\left\{-\frac{\cos \pi x/l}{1^2} + \frac{\cos 2\pi x/l}{2^2} - \frac{\cos 3\pi x/l}{3^2} + \ldots\right\}$$

$$= \frac{l^2}{3} + \frac{4l^2}{\pi^2}\sum_{n=1}^{\infty}\frac{(-1)^n}{n^2}\cos\left(\frac{n\pi x}{l}\right).$$

4.3 Any Interval $[a, b]$

If f is bounded, integrable and piecewise monotonic in $[a, b]$, then the series

$$\frac{1}{2}a_0 + \sum_{n=1}^{\infty}\left[a_n \cos\frac{n\pi(2x-a-b)}{b-a} + b_n \sin\frac{n\pi(2x-a-b)}{b-a}\right],$$

where $$a_n = \frac{2}{b-a}\int_a^b f \cos\frac{n\pi(2x-a-b)}{b-a}\,dx,$$

$$b_n = \frac{2}{b-a}\int_a^b f \sin\frac{n\pi(2x-a-b)}{b-a}\,dx$$

represents $\frac{1}{2}\left[f(x-) + f(x+)\right]$ in $a < x < b$, *and*

$$\frac{1}{2}\left[f(a+) + f(b-)\right] \text{at } x = a, b$$

and is periodic with period $(b - a)$.

On making the substitution $y = \dfrac{\pi(2x-a-b)}{b-a}$, we see that y varies from $-\pi$ to π, as x varies from a to b.

Considering the function F, where

$$f(x) = f\left(\frac{y(b-a) + \pi(a+b)}{2\pi}\right) = F(y)$$

and proceeding as above, we get the required result.

Note: The transformation is obtained by determining the two constants l and m in $y = lx + m$ such that $y = -\pi$ when $x = a$, and $y = \pi$ when $x = b$.

CHAPTER 14

4.4 The Interval [0, *l*]

If a function f satisfies the conditions of the main theorem (Dirichlet's criterion) on $[0, l]$, the substitution $y = \pi x/l$ determines a function F, where

$$f(x) = f(yl/\pi) = F(y)$$

which satisfies the conditions of the main theorem on $[0, \pi]$. The function F may now be extended [as in § 3.2] as an odd or an even function on $[-\pi, \pi]$, so as to give series consisting of terms of sines or of cosines only.

We may, if we desire, reverse the above process—that first extend it as an odd or as an even function on $[-l, l]$ and then use the substitution $y = \pi x/l$ to transform it to a function F on $[-\pi, \pi]$.

Alternatively, like § 4.3 transformation $y = \pi(2x - l)/l$ may be used to get a function F satisfying the conditions of the main theorem on $[-\pi, \pi]$ where

$$f(x) = f\left(\frac{ly + \pi l}{2\pi}\right) = F(y)$$

Example 6. Expand the periodic function $x^2, 0 \le x \le l$, of period l, in a series of (*i*) sines only, (*ii*) cosines only, (*iii*) sines and cosines, of multiples of x.

Also find the sum of the series at $x = 0, l$.

■ (*i*) The function may be extended as an odd function in $[-l, l]$, by redefining it as

$$f(x) = \begin{cases} -x^2, \text{ for } -l < x \le 0 \\ x^2, \text{ for } 0 < x \le l. \end{cases}$$

Substitution of $y = \pi x/l$ transforms it into an odd periodic function on $[-\pi, \pi]$, so that the Fourier coefficients are

$$a_n = 0 \text{ for } n = 0, 1, 2, 3, ..$$

$$b_n = \frac{1}{l}\left[\int_{-l}^{0} -x^2 \sin(n\pi x/l)\,dx + \int_0^l x^2 \sin(n\pi x/l)\,dx\right]$$

$$= \frac{2}{l}\int_0^l x^2 \sin(n\pi x/l)\,dx$$

$$= \frac{2}{l}\frac{l}{n\pi}\left[\left| -x^2 \cos\frac{n\pi x}{l}\right|_0^l + \int_0^l 2x \cos\frac{n\pi x}{l}\,dx\right]$$

$$= -\frac{2l^2}{n\pi}\cos n\pi + \frac{4l^2}{n^3\pi^3}(\cos n\pi - 1)$$

$$= \begin{cases} -\dfrac{2l^2}{n\pi}, \text{ for } n \text{ even} \\ \dfrac{2l^2}{n\pi} - \dfrac{8l^2}{n^3\pi^3} \text{ for } n \text{ odd} \end{cases}$$

The function is continuous at all points of [–*l*, *l*] except $\pm l$. Therefore, the Fourier series for all points of [0, *l*] except *l*, is

$$x^2 + \sum_{n=1}^{\infty}\left[\frac{-2(-1)^n l^2}{n\pi} - \frac{4l^2\left\{1-(-1)^n\right\}}{n^3\pi^3}\right]\sin\frac{n\pi x}{l} \qquad ...(1)$$

At $x = 0$, a point of continuity of the function, the sum of the series is zero, a fact which may be verified directly from the series.

At $x = l$, the sum of the series $= \frac{1}{2}\left[f(l-)+f(-l+)\right] = \frac{1}{2}\left(l^2 - l^2\right) = 0$

which is true on actual verification of the sum of the series.

Hence the relation (1) holds for all points of [0, *l*].

(ii) Let us now extend the function as an even function on [–*l*, *l*], by redefining it as

$$f(x) = x^2, -l \le x \le l$$

Substitution of $y = \pi x/l$ transforms *f* as an even function on $[-\pi, \pi]$, so that the Fourier coefficients are

$$b_n = 0, \text{ for } n = 1, 2, 3, \ldots$$

$$a_0 = \frac{1}{l}\int_{-l}^{l} x^2 dx = \frac{2}{l}\int_0^l x^2\, dx = \frac{2l^2}{3}$$

$$a_n = \frac{2}{l}\int_0^l x^2 \cos\frac{n\pi x}{l}\, dx = \frac{4l^2}{\pi^2 n^2}\cos n\pi = \frac{(-1)^n 4l^2}{\pi^2 n^2}$$

The function is continuous at all points of [–*l*, *l*]. Therefore for all x in [0, *l*], we have

$$x^2 = \frac{l^2}{3} + \frac{4l^2}{\pi^2}\sum_{n=1}^{\infty}\frac{(-1)^n}{n^2}\cos\frac{n\pi x}{l}$$

At $x = 0$,

$$0 = \frac{l^2}{3} - \frac{4l^2}{\pi^2}\left(1 - \frac{1}{2^2} + \frac{1}{3^2} - \ldots\right)$$

$$\therefore \qquad 1 - \frac{1}{2^2} + \frac{1}{3^2} - \frac{1}{4^2} + \ldots = \frac{\pi^2}{12}$$

At $x = l$,

$$l^2 = \frac{l^2}{3} + \frac{4l^2}{\pi^2} \sum_{n=1}^{\infty} \frac{(-1)^n}{n^2} \cos n\pi$$

$$= \frac{l^2}{3} + \frac{4l^2}{\pi^2} \Sigma \frac{1}{n^2}$$

$$\therefore\ 1 + \frac{1}{2^2} + \frac{1}{3^2} + \ldots = \frac{\pi^2}{6}$$

(*iii*) Let us apply § 4.3.

If we make the substitution $y = \pi(2x - l)/l$, then y varies from $-\pi$ to π as x varies from 0 to l. The Fourier coefficients are

$$a_0 = \frac{2}{l}\int_0^l x^2 dx = \frac{2}{3} l^2$$

$$a_n = \frac{2}{l}\int_0^l x^2 \cos\frac{n\pi(2x - l)}{l} dx = \frac{2}{n\pi}\int_0^l -x\sin\frac{n\pi(2x - l)}{l} dx$$

$$= \frac{l}{\pi^2 n^2}\left| x\cos\frac{n\pi(2x - l)}{l}\right|_0^l = \frac{l^2}{n^2}\frac{(-1)^n}{\pi^2}$$

$$b_n = \frac{2}{l}\int_0^l x^2 \sin\frac{n\pi(2x - l)}{l} dx$$

$$= -\frac{l^2}{n\pi}\cos n\pi + \frac{2}{n\pi}\int_0^l x\cos\frac{n\pi(2x - l)}{l} dx = \frac{l^2}{\pi}\frac{(-1)^{n-1}}{n}.$$

Hence for all points between 0 and l,

$$x^2 = \frac{l^2}{3} + \frac{l^2}{\pi^2}\sum_{n=1}^{\infty}\frac{(-1)^n}{n^2}\cos\frac{n\pi(2x - l)}{l} + \frac{l^2}{\pi}\sum_{n=1}^{\infty}\frac{(-1)^{n-1}}{n}\sin\frac{n\pi(2x - l)}{l}$$

At $x = 0, l$, the sum of the series $= \frac{1}{2}(l^2)$.

$$\therefore \qquad \frac{l^2}{2} = \frac{l^2}{3} + \frac{l^2}{\pi^2}\Sigma\frac{1}{n^2}$$

$$\Rightarrow \qquad 1 + \frac{1}{2^2} + \frac{1}{3^2} + \frac{1}{4^2} + \ldots = \frac{\pi^2}{6}.$$

Example 7. Expand the periodic function, of period $2l > 0$,

$$f(x) = \left|\cos\left(\frac{\pi x}{l}\right)\right|, \text{ in a Fourier series.}$$

- We have

$$f(x) = \begin{cases} -\cos(\pi x/l), & -l \le x \le -l/2, \\ \cos(\pi x/l), & -l/2 \le x \le l/2, \\ -\cos(\pi x/l), & l/2 \le x \le l \end{cases}$$

Since $f(x)$ is an even function, therefore $b_n = 0, \ \forall n$.

Now,

$$a_0 = \frac{1}{l}\int_{-l}^{l} \left|\cos(\pi x/l)\right| dx$$

$$= -\frac{1}{l}\int_{-l}^{-l/2} \cos(\pi x/l)\, dx + \frac{1}{l}\int_{-l/2}^{l/2} \cos(\pi x/l)\, dx - \frac{1}{l}\int_{l/2}^{l} \cos(\pi x/l)\, dx$$

$$= -\frac{2}{l}\int_{l/2}^{l} \cos(\pi x/l)\, dx + \frac{2}{l}\int_{0}^{l/2} \cos(\pi x/l)\, dx = 4/\pi$$

Similarly,

$$a_n = -\frac{1}{l}\int_{-l}^{-l/2} \cos(\pi x/l)\cos(n\pi x/l)\, dx + \int_{-l/2}^{l/2} \cos(\pi x/l)\cos(n\pi x/l)\, dx$$

$$-\int_{l/2}^{l} \cos(\pi x/l)\cos(n\pi x/l)\, dx$$

$$= -\frac{1}{l}\int_{l/2}^{l}\left[\cos\left\{\frac{(n+1)\pi x}{l}\right\} + \cos\left\{\frac{(n-1)\pi x}{l}\right\}\right]$$

$$+\frac{1}{l}\int_{0}^{l/2}\left[\cos\left\{\frac{(n+1)\pi x}{l}\right\} - \cos\left\{\frac{(n-1)\pi x}{l}\right\}\right] dx = -\frac{2\cos(n\pi/2)}{(n^2-1)\pi}$$

$$\therefore \quad a_{2n} = \frac{2(-1)^{n+1}}{(4n^2-1)\pi}, \ \forall n \text{ and } a_{2n-1} = 0, \ \forall n$$

Hence, $\left|\cos(\pi x/l)\right| = \dfrac{4}{\pi} + \sum\limits_{n=1}^{\infty} \dfrac{2(-1)^{n+1}}{\pi(4n^2-1)}\cos\left(\dfrac{2n\pi x}{l}\right)$.

CHAPTER 14

Example 8. Obtain the Fourier series, in the interval [–1/2, 1/2], of the function f given by

$$f(x)=\begin{cases} x-[x]-1/2, & \text{when } x \text{ is not an integer} \\ 0, & \text{otherwise} \end{cases}$$

where $[x]$ is the greatest integer $\leq x$.

■ We have,

$$f(x+1)=x+1-[x+1]-1/2=x+1-([x]+1)-1/2=x-[x]-1/2=f(x)$$

$\therefore f$ is periodic with period 1.

Also

$$f(-x)=-x-[-x]-1/2=-\{x+[-x]+1/2+1-1\}$$
$$=-\{x+[1-x]-1/2\}=-\{x-[x]-1/2\}=-f(x)$$

$\therefore f$ is an odd function of x, and so $a_n=0,\ \forall\ n$

Now

$$b_n=2\int_{-1/2}^{1/2} f(x)\sin(2n\pi x)dx$$
$$=4\int_0^{1/2} f(x)\sin(2n\pi x)dx$$
$$=4\int_0^{1/2}(x-1/2)\sin(2n\pi x)dx$$
$$=-4x\frac{\cos(2n\pi x)}{2n\pi}\Bigg|_0^{1/2}+4\int_0^{1/2}\frac{\cos(2n\pi x)}{2n\pi}dx+\frac{\cos(2n\pi x)}{n\pi}\Bigg|_0^{1/2}$$
$$=-\frac{\cos n\pi}{n\pi}+0+\frac{\cos n\pi}{n\pi}-\frac{1}{n\pi}=-\frac{1}{n\pi}.$$

Hence

$$x-[x]-\frac{1}{2}=\sum_{n=1}^{\infty}(-1/n\pi)\sin(2n\pi x).$$

Example 9. Obtain the Fourier series of the function f given by the following graph.

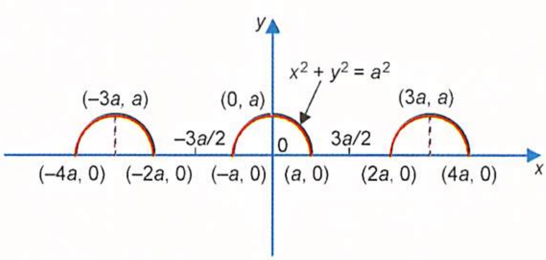


Fig. 5

- Clearly it is the graph of a continuous periodic function f of period $3a$ in $[-a, 2a]$, where

$$f(x)=\begin{cases}\sqrt{a^2-x^2}, & -a \le x \le a\\ 0, & a \le x \le 2a\end{cases}$$

Hence for all values of x, we have

$$f(x)=\frac{1}{2}a_0+\sum_{n=1}^{\infty}\left[a_n\cos\left\{n\pi(2x-a)/3a\right\}+b_n\sin\left\{n\pi(2x-a)/3a\right\}\right]$$

where
$$a_n=\frac{2}{3a}\int_{-a}^{a}\sqrt{a^2-x^2}\cos\left\{n\pi(2x-a)/3a\right\}dx$$

$$b_n=\frac{2}{3a}\int_{-a}^{a}\sqrt{a^2-x^2}\sin\left\{n\pi(2x-a)/3a\right\}dx.$$

Aliter. The graph can also be considered as that of a continuous periodic function f of period $3a$ in $\left[-\frac{3a}{2}, \frac{3a}{2}\right]$, where

$$f(x)=\begin{cases}0, & \text{for } -3a/2 \le x \le -a\\ & \text{and } a \le x \le 3a/2\\ \sqrt{a^2-x^2}, & \text{for } -a \le x \le a\end{cases}$$

It is an even function and therefore the Fourier coefficients will consist of cosine terms only. Since it is a continuous function, therefore for all values of x, we have

$$f(x)=\frac{1}{2}a_0+\sum_{n=1}^{\infty}a_n\cos(n\pi x/a),$$

where
$$a_n=\frac{1}{a}\int_{-a}^{a}\sqrt{a^2-x^2}\cos(n\pi x/a)\,dx.$$

EXERCISE

1. Expand f in a Fourier series in the interval $[-\pi, \pi]$, where $f(x) = 1$, for $-\pi < x \le 0$, and $f(x) = -2$, for $0 < x \le \pi$.
2. Obtain the Fourier series in $[-\pi, \pi]$ for the function

$$f(x)=\begin{cases}x, & \text{if } -\pi < x \le 0,\\ 2x, & \text{if } 0 \le x \le \pi\end{cases}$$

3. Are the following trigonometric series, the Fourier series?

(i) $\dfrac{\sin x}{1^2}+\dfrac{\cos 2x}{2^2}+\dfrac{\sin 3x}{3^2}+\dfrac{\cos 4x}{4^2}+\ldots,$

(ii) $\dfrac{\cos x}{\sqrt{1}}+\dfrac{\sin 2x}{\sqrt{2}}+\dfrac{\cos 3x}{\sqrt{3}}+\dfrac{\sin 4x}{\sqrt{4}}+\ldots,$

(iii) $\sum_{n=1}^{\infty} \frac{\sin nx}{\sqrt{n}}$,

(iv) $\sum_{n=1}^{\infty} \frac{\cos nx + \sin nx}{\sqrt{n}}$,

(v) $\sum_{n=1}^{\infty} \frac{\sin nx}{n^2}$.

4. Show that the Fourier series

$$\frac{\pi^2}{3} + 4 \sum_{n=1}^{\infty} (-1)^{n-1} \left(\frac{\sin nx}{2n} - \frac{\cos nx}{n^2} \right)$$

converges to the periodic function f in $]-\pi, \pi[$, where $f(x) = x^2 + x$, for $-\pi < x < \pi$, and $f(x) = \pi^2$, for $x = \pm\pi$.

What is the value of the series at

$$x = \pm\pi, 0, 9\pi, 19\pi/2, 10\pi?$$

Deduce that

$$\frac{\pi^2}{6} = 1 + \frac{1}{2^2} + \frac{1}{3^2} + \ldots.$$

5. Find the trigonometrical series which converges in $[-\pi, \pi]$ to the function

$$f(x) = \begin{cases} -\cos x, \text{ for } -\pi < x < 0, \\ \cos x, \text{ for } 0 < x < \pi. \end{cases}$$

6. Find the series of sines and cosines of multiples of x for the functions, $|\sin x|$ and $|\cos x|$ in the interval $[-\pi, \pi]$.

7. Expand the following functions in Fourier sine series:

(i) $f(x) = \begin{cases} \sin(\pi x/l), \text{ if } 0 \le x < l/2, \\ 0, \quad\quad \text{ if } l/2 < x \le l \end{cases}$

(ii) $f(x) = \begin{cases} \sin(\pi x/l), \text{ if } 0 \le x < l/2, \\ -\sin(\pi x/l), \text{ if } l/2 < x \le l. \end{cases}$

8. Show that for all values of x in $[-\pi, \pi]$, when k is not an integer,

$$\cos kx = \frac{\sin k\pi}{\pi} \left[\frac{1}{k} + \sum_{n=1}^{\infty} \frac{(-1)^n 2k \cos nx}{k^2 - n^2} \right]$$

Deduce that

$$\pi \cot k\pi = \frac{1}{k} + \sum_{n=1}^{\infty} \frac{2k}{k^2 - n^2},$$

$$\frac{\pi}{\sin k\pi} = \sum_{n=0}^{\infty} (-1)^n \left(\frac{1}{n+k} + \frac{1}{n+1-k} \right)$$

9. Find the Fourier series in $[0, \pi]$ for the function

$$f(x) = \begin{cases} \pi/3, & \text{for } 0 < x < \pi/3 \\ 0, & \text{for } \pi/3 < x < 2\pi/3 \\ -\pi/3, & \text{for } 2\pi/3 < x < \pi \end{cases}$$

Find the sum of the series when $x = \pi$, $2\pi/3$.

10. Expand the following function in a Fourier series in $[-\pi, \pi]$:

$$f(x) = \begin{cases} -\frac{1}{2}(\pi + x), & \text{when } -\pi \le x < 0 \\ \frac{1}{2}(\pi - x), & \text{when } 0 \le x < \pi \end{cases}$$

11. Expand the function $f(x) = \cos 2x$ in a series of sines in $[0, \pi]$.
12. Expand the function $e^x - 1$ in a Fourier series in $[0, 2\pi]$.
13. Expand $|x|$ in a Fourier series on $[-l, l]$.
14. Find the Fourier series of e^x in $[-l, l]$.
15. Expand the function $f(x) = 2x$ in a series of sines in $[0, 1]$.
16. Expand the function

$$f(x) = \begin{cases} x, & \text{when } 0 < x \le 1 \\ 2 - x, & \text{when } 1 < x < 2 \end{cases}$$

in the interval $[0, 2]$ as a (*i*) series of sines (*ii*) series of cosines.

17. Obtain the Fourier series for functions of the following graphs:

(*i*)

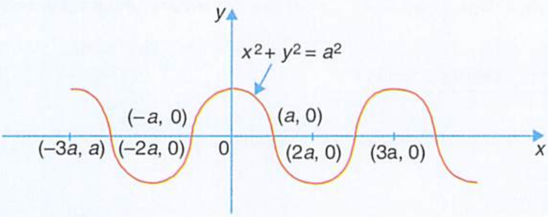


Fig. 6

(*ii*)

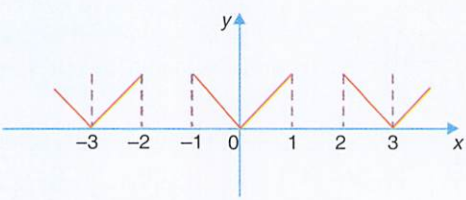


Fig. 7

[*Hint*: (*i*) Either, a continuous periodic function of period $4a$ in $[-a, 3a[$, where $f(x) = \sqrt{a^2 - x^2}$ when $-a \le x \le a$ and $f(x) = -\sqrt{a^2 - x^2}$, when $a \le x \le 3a$.

Or, a continuous periodic function of period $4a$ in $[-2a, 2a]$, where $f(x) = \sqrt{a^2 - x^2}$, for $-a \le x \le a$ and $f(x) = -\sqrt{a^2 - x^2}$ when $-2a \le x \le -a$, $a \le x \le 2a$.

(*ii*) A periodic function of period 3 in $[-3/2, 3/2]$, where $f(x) = |x|$ for $-1 < x < 1$ and $f(x) = 0$ for $-3/2 \le x \le -1, 1 \le x \le 3/2]$.

ANSWERS

1. $-\frac{1}{2} - \frac{6}{\pi}\sum_{n=1}^{\infty}\frac{\sin(2n-1)x}{2n-1}$.

2. $\frac{\pi}{4} - \frac{2}{\pi}\left(\frac{\cos x}{1^2} + \frac{\cos 3x}{3^2} + \frac{\cos 5x}{5^2} + \ldots\right) + 3\left(\frac{\sin x}{1} - \frac{\sin 2x}{2} + \frac{\sin 3x}{3} - \ldots\right)$.

3. (*i*), (*v*) Yes; (*ii*), (*iii*), (*iv*) No.

5. $\frac{4}{\pi}\left(\frac{2}{1\cdot 3}\sin 2x + \frac{4}{3\cdot 5}\sin 4x + \frac{6}{5\cdot 7}\sin 6x + \ldots\right)$.

6. $\frac{2}{\pi} - \sum_{n=1}^{\infty}\frac{4\cos(2nx)}{(4n^2-1)\pi}, \frac{2}{\pi} - \sum_{n=1}^{\infty}\frac{4(-1)^n\cos(2nx)}{\pi(4n^2-1)}$.

7. (*i*) $\frac{1}{2}\sin(\pi x/l) - \frac{4}{\pi}\sum_{n=1}^{\infty}\frac{(-1)^n n}{(4n^2-1)}\sin(2\pi nx/l), 0 \le x \le l,$ except for the value $x = l/2$, where the sum equals 1/2.

   (*ii*) $\frac{4}{\pi}\left(\frac{2}{3}\sin\frac{2\pi x}{l} - \frac{4}{15}\sin\frac{4\pi x}{l} + \frac{6}{35}\sin\frac{6\pi x}{l} - \ldots\right), 0 \le x \le l,$ except for the value $x = l/2$, where the sum equals 0.

9. $\frac{\sin 2x}{1} + \frac{\sin 4x}{2} + \frac{\sin 8x}{4} + \frac{\sin 10x}{5} + \frac{\sin 14x}{7} + \ldots$

10. $\Sigma\frac{\sin nx}{n}$.

11. $-\frac{4}{\pi}\Sigma\frac{(2n-1)\sin(2n-1)x}{2^2-(2n-1)^2}$

12. $\frac{e^{2\pi}-1}{\pi}\left[\frac{1}{2} + \Sigma\left(\frac{\cos nx}{1+n^2} - \frac{n\sin nx}{1+n^2}\right)\right] - 1$.

13. $\frac{l}{2} - \frac{4l}{\pi^2}\Sigma\frac{\cos\{(2n-1)\pi x/l\}}{(2n-1)^2}$.

14. $\frac{\sinh l}{l} + 2l\sinh l\,\Sigma\frac{(-1)^n\cos n\pi x/l}{l^2+n^2\pi^2} - 2\pi\sinh l\,\Sigma\frac{(-1)^n n\sin n\pi x/l}{l^2+n^2\pi^2}$.

15. $\frac{4}{\pi}\Sigma(-1)^{n+1}\frac{\sin n\pi x}{n}$.

16. (*i*) $\dfrac{8}{\pi^2}\Sigma(-1)^{n-1}\dfrac{\sin\{(2n-1)\pi x/2\}}{(2n-1)^2}$,

(*ii*) $\dfrac{1}{2}-\dfrac{4}{\pi^2}\Sigma\dfrac{\cos(2n-1)\pi x}{(2n-1)^2}$.

17. (*i*) $f(x)=\dfrac{1}{2}a_0+\Sigma\left[a_n\cos\{nx(x-a)/2a\}+b_n\sin\{nx(x-a)/2a\}\right]$, where

$$a_n=\frac{1}{2a}\int_{-a}^{a}\sqrt{a^2-x^2}\cos\{n\pi(x-a)/2a\}\,dx-\frac{1}{2a}\int_{a}^{3a}\sqrt{a^2-x^2}\cos\{n\pi(x-a)/2a\}\,dx$$

and similar expression for b_n with cosines replaced by sines.

(*ii*) $\dfrac{1}{3}+\dfrac{2}{\pi}\Sigma\left[\dfrac{\sin 2\pi n/3}{n}-\dfrac{3(1-\cos 2\pi n/3)}{2\pi n^2}\right]\cos\dfrac{2\pi nx}{3}$

$$=\frac{1}{3}+\frac{\sqrt{3}}{\pi}\left[\frac{\cos 2\pi x/3}{1}-\frac{\cos 4\pi x/3}{2}+\frac{\cos 8\pi x/3}{4}-\dots\right]$$

$$-\frac{9}{2\pi^2}\left[\frac{\cos 2\pi x/3}{1^2}+\frac{\cos 4\pi x/3}{2^2}+\frac{\cos 8\pi x/3}{4^2}+\dots\right].$$

15 Functions of Several Variables

So far attention has mainly been directed to functions of a single independent variable and the application of the differential calculus to such functions has been considered. In this chapter, we shall be mainly concerned with the application of differential calculus to functions of more than one variable. The characteristic properties of a function of n independent variables may usually be understood by the study of a function of two or three variables and this restriction of two or three variables will be generally maintained. This restriction has the considerable advantage of simplifying the formulae and of reducing the mechanical labour.

1. EXPLICIT AND IMPLICIT FUNCTIONS

If we consider a set of n independent variables $x, y, z, ..., t$ and one dependent variable u, the equation

$$u = f(x, y, z, ..., t) \qquad ...(1)$$

denotes the functional relation. In this case if $x_1, y_1, z_1, ..., t_1$, are the n arbitrarily assigned values of the independent variables, the corresponding values of the dependent variable u are determined by the function relation.

The function represented by equation (1) is an *explicit* function but where several variables are concerned it is rarely possible to obtain an equation expressing one of the variables explicitly in terms of the others. Thus most of the functions of more than one variable are *implicit* functions, that is to say we are given a functional relation

$$\phi(x, y, z, ..., t) = 0 \qquad ...(2)$$

connecting the n variables $x, y, z, ..., t$, and is not in general possible to solve this equation to find an *explicit* function which expresses one of these variables say x, in terms of the other $n-1$ variables.

In this chapter we shall be mainly concerned with the *explicit* functions.

1.1 An Explicit Function of Two Variables

If x, y are two independent variables and a variable z depends for its values on the values of x, y by a functional relation

$$z = f(x, y) \qquad ...(3)$$

then we say z is a *function of* x, y. The ordered pair of numbers (x, y) is called a *point* and the aggregate of the pairs of numbers (x, y) is said to be the *domain* (or region) *of definition* of the function.

When the domain of definition is bounded by a closed curve C, it is said to be *closed* if f is defined for all points within and on the curve C; but *open or unclosed* when the function is defined for points within but not on the curve C.

1.2 The Neighbourhood of a Point

The set of values x_1, y_1 other than a, b that satisfy the conditions

$$|x_1 - a| < \delta, |y_1 - b| < \delta$$

where δ is an arbitrarily small positive number, is said to form a *neighbourhood* of the point (a, b). Thus a neighbourhood is the square

$$(a - \delta, a + \delta; b - \delta, b + \delta)$$

where x takes any value from $a - \delta$ to $a + \delta$ except a, and y from $b - \delta$ to $b + \delta$ except b.

This is not the only way of specifying a neighbourhood of a point. There can be many other, though equivalent ways; for example the points inside the circle $x^2 + y^2 = \delta^2$ may be taken as a neighbourhood of the point $(0, 0)$.

1.3 Limit Point

A point (ξ, η) is called a *limit point* or *a point of condensation* of a set of points S, if for every neighbourhood of (ξ, η) contains an infinite number of points of S. The limit point itself may or may not be a point of the set. For example, the point $(0, 0)$ is a limit point of the set $\{(1/m, 1/n) : m, n \in \mathbf{N}\}$.

1.4 The Limit of a Function

A function f is said to tend to a limit l as a point (x, y) tends to the point (a, b) if for every arbitrarily small positive number ε, there corresponds a positive number δ, such that

$$|f(x, y) - l| < \varepsilon,$$

for every point (x, y), [different from (a, b)] which satisfies

$$|x - a| < \delta, |y - b| < \delta$$

In other words, *a function f tends to a limit l, when (x, y) tends to (a, b) if for every positive number ε, there corresponds a neighbourhood N of (a, b) such that*

$$|f(x, y) - l| < \varepsilon,$$

for every point (x, y) other than (a, b) of the neighbourhood N.

Symbolically we then write

$$\lim_{(x, y)\to(a, b)} f(x, y) = l.$$

l is the *limit* (the *double limit* or the *simultaneous limit*) of f when x, y tend to a & b simultaneously.

Remark: The above definition implies that there must be no assumption of any relation between the independent variables as they tend to their respective limits.

For instance take $f(x, y)$ where

$$f(x, y) = \frac{xy}{x^2 + y^2}$$

and find the limit when $(x, y) \to (0, 0)$.

If we put $y = m_1 x$ and let $x \to 0$, we get the limit to be equal to $\dfrac{m_1}{1+m_1^2}$, while putting $y = m_2 x$ leads to a limit $\dfrac{m_2}{1+m_2^2}$. Similarly letting $x \to 0$, while y remains constant or vice-versa leads to zero limit. Thus, we are led to erroneous results. Geometrically speaking when we approach the point (0, 0) along different paths, first along lines with slopes m_1 and m_2 and then along lines parallel to the coordinate axes, the function reaches different limits. The simultaneous limit postulates that by whatever path the point is approached, the function f attains the same limit. In general the determination whether a simultaneous limit exists or not is a difficult matter but very often a simple consideration enables us to show that the *limit does not exist*.

It may however be noted that

$$\lim_{(x,\,y)\to(a,\,b)} f(x, y) = l \;\Rightarrow\; \lim_{x\to a} f(x, b) = l = \lim_{y\to b} f(a, y)$$

Non-existence of limit. The above remark makes it amply clear that if $\lim\limits_{(x,\,y)\to(a,\,b)} f(x, y) = l$ and if $y = \phi(x)$ is any function such that $\phi(x) \to b$, when $x \to a$, then $\lim\limits_{x\to a} f(x, \phi(x))$ must exist and should be equal to l.

Thus, if we can find two functions $\phi_1(x)$ and $\phi_2(x)$ *such that the limits of* $f(x, \phi_1(x))$ *and* $f(x, \phi_2(x))$ *are different, then the simultaneous limit in question does not exist.*

Example 1(*a*). Let

$$f(x, y) = \begin{cases} \dfrac{x^2 y}{x^4 + y^2}, & \text{if } \; x^4 + y^2 \neq 0 \\ 0, & \text{if } \; x + y = 0 \end{cases}$$

If we approach the origin along any axis, $f(x, y) \equiv 0$.

If we approach (0, 0) along any line $y = mx$, then

$$f(x, y) = f(x, mx) = \frac{mx^3}{x^4 + m^2x^2} = \frac{mx}{x^2 + m^2} \to 0, \text{ as } x \to 0$$

So any straight line approach gives,

$$\lim_{(x,\,y)\to(0,\,0)} f(x, y) = 0$$

But putting $y = mx^2$,

$$\lim_{(x,\,y)\to(0,\,0)} f(x, y) = \lim_{x\to 0} f(x, mx^2) = \frac{m}{1+m^2}$$

which is different for the different m selected.

Hence, $\lim\limits_{(x,\,y)\to(0,\,0)} f(x, y)$ does not exist.

Thus, the function possesses no limit at the origin, but a straight line approach gives the limit zero.

Example 1 (*b*). Show that

$$\lim_{(x,\,y)\to(0,\,0)} \frac{2xy^2}{x^2+y^4} \text{ does not exist.}$$

- If we put $x = my^2$ and let $y \to 0$, we get

$$\lim_{y\to 0} \frac{2my^4}{(m^2+1)y^4} = \frac{2m}{1+m^2}$$

which is different for different values of m.

Hence, the limit does not exist.

Remark: It is pointed out earlier also that the determination of a simultaneous limit is a difficult matter but a simple consideration, as shown above, very often, enables us to show that the limit does not exist. We now show that sometimes it is possible to determine the simultaneous limit by changing to polars.

Example 2 (*a*). Show that

$$\lim_{(x,\,y)\to(0,\,0)} xy\frac{x^2-y^2}{x^2+y^2} = 0$$

- Put $x = r\cos\theta$, $y = r\sin\theta$

$$\left| xy\frac{x^2-y^2}{x^2+y^2} \right| = \left| r^2 \sin\theta \cos\theta \cos 2\theta \right|$$

$$= \left| \frac{r^2}{4} \sin 4\theta \right| \le \frac{r^2}{4} = \frac{x^2+y^2}{4} < \varepsilon,$$

if

$$\frac{x^2}{4} < \frac{\varepsilon}{2}, \frac{y^2}{4} < \frac{\varepsilon}{2}$$

or if

$$|x| < \sqrt{2\varepsilon} = \delta, \ |y| < \sqrt{2\varepsilon} = \delta$$

Thus for $\varepsilon > 0$, $\exists\, \delta > 0$ such that

$$\left| xy\frac{x^2-y^2}{x^2+y^2} - 0 \right| < \varepsilon, \text{ when } |x| < \delta, |y| < \delta$$

$$\Rightarrow \qquad \lim_{(x,\,y)\to(0,\,0)} xy\frac{x^2-y^2}{x^2+y^2} = 0$$

Example 2 (*b*). Show that

$$\lim_{(x,\, y)\to(0,\, 0)} \frac{\sqrt{x^2y^2+1}-1}{x^2+y^2} = 0$$

■ Since x, y are small

$$\frac{\sqrt{x^2y^2+1}-1}{x^2+y^2} = \frac{(1+x^2y^2)^{1/2}-1}{x^2+y^2} \simeq \frac{\frac{1}{2}x^2y^2}{x^2+y^2}$$

Now changing to polars, we can show, as in the above example, that

$$\lim_{(x,\, y)\to(0,\, 0)} \frac{\frac{1}{2}x^2y^2}{x^2+y^2} = 0$$

Hence the required result.

Ex. 1. Show that

(*i*) $\lim_{(x,\, y)\to(0,\, 0)} \left(\frac{1}{|x|}+\frac{1}{|y|}\right) = \infty$, (*ii*) $\lim_{(x,\, y)\to(0,\, 0)} \frac{xy^2}{x^2+y^2} = 0$,

(*iii*) $\lim_{(x,\, y)\to(0,\, 0)} (x+y) = 0$, (*iv*) $\lim_{(x,\, y)\to(0,\, 0)} (1/xy)\sin(x^2y+xy^2) = 0$

Ex. 2. Show that the limit, when $(x, y) \to (0, 0)$ does not exist in each case

(*i*) $\lim \frac{2xy}{x^2+y^2}$, (*ii*) $\lim \frac{xy^3}{x^2+y^6}$,

(*iii*) $\lim \frac{x^2y^2}{x^2y^2+(x^2-y^2)^2}$, (*iv*) $\lim \frac{x^3+y^3}{x-y}$

[*Hint*: (*iv*) Put $y = x - mx^3$].

Ex. 3. Show that the limit, when $(x, y) \to (0, 0)$ exist in each case.

(*i*) $\lim \frac{xy}{\sqrt{x^2+y^2}}$, (*ii*) $\lim \frac{x^3y^3}{x^2+y^2}$,

(*iii*) $\lim \frac{x^3-y^3}{x^2+y^2}$, (*iv*) $\lim \frac{x^4+y^4}{x^2+y^2}$.

1.5 Algebra of Limits

If f and g are two functions with a domain N, we define four functions, $f \pm g$, fg, f/g on N by setting

$$(f+g)(x, y) = f(x, y) + g(x, y)$$

$$(f - g)(x, y) = f(x, y) - g(x, y)$$

$$f \cdot g\,(x, y) = f(x, y) \cdot g(x, y)$$

$$(f/g)(x, y) = f(x, y)/g(x, y), \text{ if } g(x, y) \neq 0, \text{ for } (x, y) \in N.$$

Theorem 1. *If f, g be two functions defined on some neighbourhood of a point* (a, b) *such that* $\lim f(x, y) = l$, $\lim g(x, y) = m$, *when* $(x, y) \to (a, b)$, *then*

(*i*) $\lim(f + g) = \lim f + \lim g = l + m$

(*ii*) $\lim(f - g) = \lim f - \lim g = l - m$

(*iii*) $\lim(f \cdot g) = \lim f \cdot \lim g = l \cdot m$

(*iv*) $\lim \dfrac{f}{g} = \dfrac{\lim f}{\lim g} = \dfrac{l}{m}$, *provided* $m \neq 0$, *when* $(x, y) \to (a, b)$

The proofs are exactly similar to those of the corresponding theorems for a single variable.

Example 3 (*a*). Prove that

$$\lim_{(x, y) \to (1, 2)} (x^2 + 2y) = 5$$

■ *Method* 1. (Using definition of limit). We have to show that for any $\varepsilon > 0$, we can find $\delta > 0$, such that

$$|x^2 + 2y - 5| < \varepsilon, \textit{ when} |x - 1| < \delta, |y - 2| < \delta$$

If $|x - 1| < \delta$, and $|y - 2| < \delta$, then

$$1 - \delta < x < 1 + \delta \text{ and } 2 - \delta < y < 2 + \delta, \text{ excluding } x = 1, y = 2$$

Thus

$$1 - 2\delta + \delta^2 < x^2 < 1 + 2\delta + \delta^2$$

and

$$4 - 2\delta < 2y < 4 + 2\delta$$

Adding

$$5 - 4\delta + \delta^2 < x^2 + 2y < 5 + 4\delta + \delta^2$$

or

$$-4\delta + \delta^2 < x^2 + 2y - 5 < 4\delta + \delta^2$$

Now if $\delta \leq 1$, it follows that

$$-5\delta < x^2 + 2y - 5 < 5\delta$$

i.e.,

$$|x^2 + 2y - 5| < 5\delta = \varepsilon$$

so that $\delta = \varepsilon/5$ (or $\delta = 1$ whichever is smaller).

$$\therefore \qquad |x^2 + 2y - 5| < \varepsilon \text{ when } |x - 1| < \delta, |y - 2| < \delta$$

$$\therefore \qquad \lim_{(x,\, y)\to(1,\, 2)} (x^2 + 2y) = 5$$

Method 2. Using above theorem on algebra of limits,

$$\lim_{(x,\, y)\to(1,\, 2)} (x^2 + 2y) = \lim_{(x,\, y)\to(1,\, 2)} x^2 + \lim_{(x,\, y)\to(1,\, 2)} 2y = 1 + 4 = 5.$$

Example 3 (*b*). Show that

$$(i) \quad \lim_{(x,\, y)\to(0,\, 0)} \frac{x \sin (x^2 + y^2)}{x^2 + y^2} = 0, \quad (ii) \quad \lim_{(x,\, y)\to(2,\, 1)} \frac{\sin^{-1}(xy - 2)}{\tan^{-1}(3xy - 6)} = \frac{1}{3}.$$

■ (i) $\displaystyle\lim_{(x,\, y)\to(0,\, 0)} \frac{x \sin (x^2 + y^2)}{x^2 + y^2} = \lim_{(x,\, y)\to(0,\, 0)} x \cdot \lim_{(x,\, y)\to(0,\, 0)} \frac{\sin (x^2 + y^2)}{x^2 + y^2} = 0.1 = 0$

(ii) $\displaystyle\lim_{(x,\, y)\to(2,\, 1)} \frac{\sin^{-1}(xy - 2)}{\tan^{-1}(3xy - 6)} = \lim_{t\to 0} \frac{\sin^{-1} t}{\tan^{-1} 3t}$, where $t = xy - 2 = \displaystyle\lim_{t\to 0} \frac{1/\sqrt{1 - t^2}}{3/(1 + 9t^2)} = \frac{1}{3}$

Ex. 1. Show that $\displaystyle\lim_{(x,\, y)\to(0,\, 1)} \tan^{-1}(y/x)$, does not exist.

$$\left[\textit{Hint}: \text{ Limit from the left is } -\frac{\pi}{2} \text{ and that from the right } \frac{\pi}{2} \right].$$

Ex. 2. Show, by using the definition that

$$\lim_{(x,\, y)\to(1,\, 2)} 3xy = 6$$

Ex. 3. Prove that

$$(i) \quad \lim_{(x,\, y)\to(4,\, \pi)} x^2 \sin \frac{y}{8} = 8\sqrt{2}, \qquad (ii) \quad \lim_{(x,\, y)\to(0,\, 1)} e^{-1/x^2 (y - 1)^2} = 0,$$

$$(iii) \quad \lim_{(x,\, y)\to(0,\, 1)} \frac{x + y - 1}{\sqrt{x} - \sqrt{1 - y}} = 0, \; x > 0, \; y < 1.$$

1.6 Repeated Limits

If a function f is defined in some neighbourhood of (a, b), then the limit

$$\lim_{y\to b} f(x, y),$$

if it exists, is a function of x, say $\phi(x)$. If then the limit $\displaystyle\lim_{x\to a} \phi(x)$ exists and is equal to λ, we write

$$\lim_{x \to a} \lim_{y \to b} f(x, y) = \lambda$$

and say that λ is a *repeated limit* of f as $y \to b$, $x \to a$.

If we change the order of taking the limits, we get the other repeated limit

$$\lim_{y \to b} \lim_{x \to a} f(x, y) = \lambda' \text{ (say)}$$

when first $x \to a$, and then $y \to b$.

These two limits may or may not be equal.

Note: In case the simultaneous limit exists, these two repeated limits if they exist are necessarily equal but the converse is not true. However *if the repeated limits are not equal, the simultaneous limit cannot exist.*

Example 4. (*i*) Let

$$f(x, y) = \frac{xy}{x^2 + y^2}, \text{ then}$$

$$\lim_{y \to 0} \lim_{x \to 0} f(x, y) = \lim_{y \to 0} (0) = 0,$$

$$\lim_{x \to 0} \lim_{y \to 0} f(x, y) = 0.$$

Thus, the repeated limits exist and are equal. But the simultaneous limit does not exist which may be seen by putting $y = mx$.

(*ii*) Let

$$f(x, y) = \frac{y - x}{y + x} \cdot \frac{1 + x}{1 + y}, \text{ then}$$

$$\lim_{x \to 0} \lim_{y \to 0} f(x, y) = \lim_{x \to 0} \left(-\frac{1 + x}{1} \right) = -1,$$

$$\lim_{y \to 0} \lim_{x \to 0} f(x, y) = \lim_{y \to 0} \left(\frac{1}{1 + y} \right) = 1.$$

Thus, the two repeated limits exist but are unequal, consequently the simultaneous limit cannot exist, which may be verified by putting $y = mx$.

Example 5. Show that the limit exists at the origin but the repeated limits do not, where

$$f(x, y) = \begin{cases} x \sin\left(\dfrac{1}{y}\right) + y \sin\left(\dfrac{1}{x}\right), & xy \neq 0 \\ 0, & xy = 0 \end{cases}$$

■ Here $\lim_{y \to 0} f(x, y)$, $\lim_{x \to 0} f(x, y)$ do not exist and therefore $\lim_{x \to 0} \lim_{y \to 0} f(x, y)$; $\lim_{y \to 0} \lim_{x \to 0} f(x, y)$ do not exist.

Again

$$\left| x \sin \frac{1}{y} + y \sin \frac{1}{x} \right| < |x| + |y| \le 2(x^2 + y^2)^{1/2} < \varepsilon,$$

if

$$x^2 < \frac{\varepsilon^2}{4},\ y^2 < \frac{\varepsilon^2}{4}$$

or

$$|x| < \frac{\varepsilon}{2} = \delta,\ |y| < \frac{\varepsilon}{2} = \delta$$

Thus for $\varepsilon > 0, \exists\, \delta > 0$ such that

$$\left| x \sin \frac{1}{y} + y \sin \frac{1}{x} \right| < \varepsilon, \text{ when } |x| < \delta,\ |y| < \delta$$

$$\Rightarrow \qquad \lim_{(x,\, y) \to (0,\, 0)} \left(x \sin \frac{1}{y} + y \sin \frac{1}{x} \right) = 0.$$

Example 6. Show that the repeated limits exist at the origin and are equal but the simultaneous limit does not exist, where

$$f(x, y) = \begin{cases} 1, & \text{if } xy \ne 0 \\ 0, & \text{if } xy = 0 \end{cases}$$

■ Here

$$\lim_{y \to 0} f(x, y) = \begin{cases} 1, & \text{if } x \ne 0 \\ 0, & \text{if } x = 0 \end{cases}$$

$$\therefore \qquad \lim_{x \to 0} \lim_{y \to 0} f(x, y) = 1$$

Similarly,

$$\lim_{y \to 0} \lim_{x \to 0} f(x, y) = 1$$

Hence, the repeated limits exist and are equal.

Again, since there are points arbitrarily near (0, 0) at which f is equal to 0 and points arbitrarily near (0, 0) at which f is equal to 1, therefore, there is an $\varepsilon > 0$, such that

$$|f(x, y) - f(0, 0)| = |f(x, y)| \nless \varepsilon,$$

for all points in any neighbourhood of (0, 0).

Hence, $\lim_{(x,\, y) \to (0,\, 0)} f(x, y)$ does not exist.

Ex. 1. Show that $\lim\limits_{(x, y)\to(0, 0)} f(x, y)$ and $\lim\limits_{y\to 0}\lim\limits_{x\to 0} f(x, y)$ exist, but $\lim\limits_{x\to 0}\lim\limits_{y\to 0} f(x, y)$ does not, where

$$f(x, y) = \begin{cases} y + x \sin\left(\dfrac{1}{y}\right), & \text{if } y \neq 0 \\ 0 & , \text{ if } y = 0. \end{cases}$$

Ex. 2. Show that $\lim\limits_{x\to 0}\lim\limits_{y\to 0} f(x, y)$ exists, but the other repeated limit and the double limit do not exist at the origin, when

$$f(x, y) = \begin{cases} y \sin(1/x) + xy/(x^2 + y^2), & x \neq 0 \\ 0 & , x = 0. \end{cases}$$

Ex. 3. Show that the repeated limits exist but the double limit does not when $(x, y) \to (0, 0)$:

(*i*) $f(x, y) = \dfrac{x - y}{x + y}$, (*ii*) $f(x, y) = \dfrac{x^2 y^2}{x^4 + y^4 - x^2 y^2}$

(*iii*) $f(x, y) = \begin{cases} \dfrac{x^3 + y^3}{x - y}, & x \neq y \\ 0, & x = y \end{cases}$ (*iv*) $f(x, y) = \begin{cases} \dfrac{x^2 - y^2}{x^2 + y^2}, & x \neq y \\ 0, & x = y \end{cases}$

Ex. 4. Show that the limit and the repeated limits exist when $(x, y) \to (0, 0)$:

$$f(x, y) = \begin{cases} xy\,\dfrac{x^2 - y^2}{x^2 + y^2}, & (x, y) \neq (0, 0) \\ 0, & (x, y) = (0, 0). \end{cases}$$

2. CONTINUITY

A function f is said to be *continuous* at a point (a, b) of its domain of definition, if

$$\lim_{(x, y)\to(a, b)} f(x, y) = f(a, b)$$

In other words, a function f is said to be *continuous* at a point (a, b) of its domain of definition if for $\varepsilon > 0$, there exists a neighbourhood N of (a, b) such that

$$|f(x, y) - f(a, b)| < \varepsilon, \text{ for all } (x, y) \in N$$

Note: The definition of continuity of a function f at a point (a, b) requires that besides (a, b), f is defined in a certain neighbourhood of (a, b) and moreover the limit of f when $(x, y) \to (a, b)$ exists and equals to the value $f(a, b)$.

A function which is not continuous at a point is said to be *discontinuous* there at.

Remark: A point to be particularly noticed is that if a function of more than one variable is continuous at a point, it is continuous at that point when considered as a function of a single variable. To be more specific if a function f of two variables x, y is continuous at (a, b) then $f(x, b)$ is a continuous function of x at $x = a$ and $f(a, y)$ that of y at $y = b$.

The converse however is not true, *i.e.*, a function may be a continuous function of one variable when the others remain constant and yet not be a continuous function of all the variables.

For instance, consider a function f, where

$$f(x, y) = \begin{cases} \dfrac{2xy}{x^2 + y}, & (x, y) \neq (0, 0) \\ 0, & \text{at } (0, 0) \end{cases}$$

The function is not continuous at (0, 0) for $\lim\limits_{(x, y)\to(0, 0)} f(x, y)$ does not exist. But

$$\lim_{x\to 0} f(x, 0) = 0 = f(0, 0), \text{ and } \lim_{y\to 0} f(0, y) = 0 = f(0, 0)$$

so that f is continuous at (0, 0), when considered as a function of a single variable x or that of y.

A function is said to be continuous in a region if it is continuous at every point of the same.

As in limits, it can be easily proved that the sum, difference, product and quotient (provided the denominator does not vanish) of two continuous functions are also continuous.

The theorems on continuity for functions of a single variable can be easily extended to functions of several variables; the proofs for some of them, except for verbal changes, are the same while for others the method is not quite the same. However, within the scope of the present work, it is not possible to discuss all of them here.

Example 7. Investigate the continuity at (0, 0) of

$$f(x, y) = \begin{cases} \dfrac{x^2 - y^2}{x^2 + y^2}, & (x, y) \neq (0, 0) \\ 0, & (x, y) = (0, 0) \end{cases}$$

■ Since $\lim\limits_{(x, y)\to(0, 0)} f(x, y)$ does not exist, therefore the function is not continuous at (0, 0).

Example 8. Investigate for continuity at (1, 2)

$$f(x, y) = \begin{cases} x^2 + 2y, & (x, y) \neq (1, 2) \\ 0, & (x, y) = (1, 2) \end{cases}$$

■ Here

$$\lim_{(x, y)\to(1, 2)} f(x, y) = 5 \neq f(1, 2).$$

Hence, the function is not continuous at (1, 2).

The point (1, 2) is a *point of discontinuity* of the function.

However, if the function has the value 5 at (1, 2), it was then continuous at the point.

Remark: If, as in the above example, it is possible to so redefine the value of the function at a point of discontinuity that the new function is continuous, we say that the point is a *removable discontinuity* of the original function.

Example 9. Show that the function

$$f(x, y) = \begin{cases} \dfrac{xy}{\sqrt{x^2 + y^2}}, & (x, y) \neq (0, 0) \\ 0, & (x, y) = (0, 0) \end{cases}$$

is continuous at the origin.

■ Let $x = r\cos\theta$, $y = r\sin\theta$.

$\therefore$
$$\left|\frac{xy}{\sqrt{x^2 + y^2}}\right| = r\,|\cos\theta\,\sin\theta| \leq r = \sqrt{x^2 + y^2} < \varepsilon,$$

if

$$x^2 < \frac{\varepsilon^2}{2},\quad y^2 < \frac{\varepsilon^2}{2}$$

or, if

$$|x| < \frac{\varepsilon}{\sqrt{2}},\quad |y| < \frac{\varepsilon}{\sqrt{2}}$$

Thus

$$\left|\frac{xy}{\sqrt{x^2 + y^2}} - 0\right| < \varepsilon, \text{ when } |x| < \frac{\varepsilon}{\sqrt{2}}, |y| < \frac{\varepsilon}{\sqrt{2}}$$

$\Rightarrow$
$$\lim_{(x, y)\to(0, 0)} \frac{xy}{\sqrt{x^2 + y^2}} = 0$$

$\therefore$
$$\lim_{(x, y)\to(0, 0)} f(x, y) = f(0, 0)$$

Hence, f is continuous at (0, 0).

EXERCISE

1. Show that the following functions are discontinuous at the origin:

   (*i*) $f(x, y) = \begin{cases} \dfrac{1}{x^2 + y^2}, & (x, y) \neq (0, 0) \\ 0, & (x, y) = (0, 0) \end{cases}$

   (*ii*) $f(x, y) = \dfrac{x^4 - y^4}{x^4 + y^4}$, $(x, y) \neq (0, 0)$, $f(0, 0) = 0$

   (*iii*) $f(x, y) = \dfrac{(x^2 y^2)}{(x^4 + y^4)}$, $(x, y) \neq (0, 0)$, $f(0, 0) = 0$

2. Show that the following functions are continuous at the origin:

(i) $f(x, y) = \dfrac{x^2 y^2}{(x^2 + y^2)}$, $(x, y) \neq (0, 0)$, $f(0, 0) = 0$.

(ii) $f(x, y) = \begin{cases} \dfrac{x^3 y^3}{x^2 + y^2}, & (x, y) \neq (0, 0) \\ 0, & (x, y) = (0, 0). \end{cases}$

3. Show that the following functions are discontinuous at (0, 0).

(i) $f(x, y) = \begin{cases} \dfrac{x^2 y}{x^3 + y^3}, & (x, y) \neq (0, 0) \\ 0, & (x, y) = (0, 0) \end{cases}$

(ii) $f(x, y) = \begin{cases} \dfrac{x^3 + y^3}{x - y}, & x \neq y \\ 0, & x = y \end{cases}$

(iii) $f(x, y) = \dfrac{xy^3}{x^2 + y^6}$, $(x, y) \neq (0, 0)$, $f(0, 0) = 0$.

4. Discuss the following functions for continuity at (0, 0).

(i) $f(x, y) = \begin{cases} \dfrac{x^2 y}{x^3 + y^3}, & x^2 + y^2 \neq 0 \\ 0, & x + y = 0 \end{cases}$

(ii) $f(x, y) = \begin{cases} 2xy\dfrac{x^2 - y^2}{x^2 + y^2}, & (x, y) \neq (0, 0) \\ 0, & (x, y) = (0, 0) \end{cases}$

(iii) $f(x, y) = \begin{cases} 0, & (x, y) = (2y, y) \\ \exp\{|x - 2y|/(x^2 - 4xy + 4y^2)\}, & (x, y) \neq (2y, y). \end{cases}$

5. Show that f has a removable discontinuity at (2, 3):

$$f(x, y) = \begin{cases} 3xy, & (x, y) \neq (2, 3) \\ 6, & (x, y) = (2, 3) \end{cases}$$

Suitably redefine the function to make it continuous.

6. Show that the function f is continuous at the origin, where

$$f(x, y) = \begin{cases} \dfrac{x^3 - y^3}{x^2 + y^2}, & (x, y) \neq (0, 0) \\ 0, & (x, y) = (0, 0). \end{cases}$$

7. Can the given functions be appropriately defined at (0, 0) in order to be continuous there?

(i) $f(x, y) = |x|^y$, (ii) $f(x, y) = \sin\frac{x}{y}$,

(iii) $f(x, y) = \frac{x^3 + y^3}{x^2 + y^2}$, (iv) $f(x, y) = x^2 \log(x^2 + y^2)$.

3. PARTIAL DERIVATIVES

The ordinary derivative of a function of several variables with respect to one of the independent variables, keeping all other independent variables constant is called the *partial derivative* of the function with respect to the variable. Partial derivative of $f(x, y)$ with respect to x is generally denoted by $\partial f/\partial x$ or f_x or $f_x(x, y)$, while those with respect to y are denoted by $\partial f/\partial y$ or f_y or $f_y(x, y)$.

$$\therefore \quad \frac{\partial f}{\partial x} = \lim_{\delta x \to 0} \frac{f(x + \delta x, y) - f(x, y)}{\delta x}$$

and

$$\frac{\partial f}{\partial y} = \lim_{\delta y \to 0} \frac{f(x, y + \delta y) - f(x, y)}{\delta y}$$

when these limits exist.

The partial derivatives at a particular point (a, b) are often denoted by

$$\left[\frac{\partial f}{\partial x}\right]_{(a,\, b)}, \quad \frac{\partial f(a, b)}{\partial x} \quad \text{or} \quad f_x(a, b)$$

and

$$\left[\frac{\partial f}{\partial y}\right]_{(a,\, b)}, \quad \frac{\partial f(a, b)}{\partial y} \quad \text{or} \quad f_y(a, b)$$

$$\therefore \quad f_x(a, b) = \lim_{h \to 0} \frac{f(a + h, b) - f(a, b)}{h}$$

$$f_y(a, b) = \lim_{k \to 0} \frac{f(a, b + k) - f(a, b)}{k}$$

in case the limit exists.

Example 10. If $f(x, y) = 2x^2 - xy + 2y^2$, then find $\partial f/\partial x$ and $\partial f/\partial y$ at the point (1, 2).

■ Now

$$\frac{\partial f}{\partial x} = 4x - y = 2, \text{ at } (1, 2)$$

$$\frac{\partial f}{\partial y} = -x + 4y = 7, \text{ at } (1, 2)$$

Note: $f_x(1, 2)$ and $f_y(1, 2)$ have been respectively obtained from $f_x(x, y)$ and $f_y(x, y)$ by replacing (x, y) by $(1, 2)$. The procedure, though simple, is not always possible. The reader has to be on his guard.

Example 11. If

$$f(x, y) = \begin{cases} \dfrac{xy}{x^2 + y^2}, & (x, y) \neq (0, 0) \\ 0, & (x, y) = (0, 0) \end{cases}$$

show that both the partial derivatives exist at (0, 0) but the function is not continuous there at.

■ Putting $y = mx$, we see that

$$\lim_{x \to 0} f(x, y) = \frac{m}{1 + m^2}$$

so that the limit depends on the value of m, *i.e.*, on the path of approach and is different for the different paths followed and therefore does not exist. Hence the function $f(x, y)$ is not continuous at (0, 0). Again

$$f_x(0, 0) = \lim_{h \to 0} \frac{f(0 + h, 0) - f(0, 0)}{h} = \lim_{h \to 0} \frac{0}{h} = 0$$

$$f_y(0, 0) = \lim_{k \to 0} \frac{f(0 \cdot 0 + k) - f(0, 0)}{k} = \lim_{k \to 0} \frac{0}{k} = 0 .$$

Notes:

1. Unlike the situation for functions of one variable, the existence of the first partial derivatives at a point does not imply continuity at the point. The explanation lies in the fact that the information given by the existence of the two first partial derivatives at a point is limited. The values of f_x and f_y at a point (a, b) depend only on the values of f along two lines through (a, b) respectively parallel to the coordinate axes. This information is incomplete and tells us nothing at all about the behaviour of the function f as the point (a, b) is approached along lines not parallel to the axes. On the other hand, the continuity of f at (a, b) requires the function to tend to its value $f(a, b)$ by whatever path the point (a, b) is approached. Therefore, there is nothing surprising in the fact that the *partial derivatives may exist at a point at which the function is not even continuous.*

2. In the above example, if $x \neq y$,

$$f_x = \frac{y^3 - x^2 y}{(x^2 + y^2)^2},$$

$$f_y = \frac{x^3 - xy^2}{(x^2 + y^2)^2},$$

and $f_x(0, 0)$, $f_y(0, 0)$ cannot be computed from them by letting $x = 0, y = 0$.

EXERCISE

1. If $f(x, y) = x^3 y + e^{xy^2}$, find f_x and f_y.

2. If $f(x, y) = xy\dfrac{(x^2 - y^2)}{(x^2 + y^2)}$, when $x^2 + y^2 \neq 0$, and $f(0, 0) = 0$, show that

$$f_x(x, 0) = 0 = f_y(0, y)$$

$$f_x(0, y) = -y, f_y(x, 0) = x.$$

3. If $f(x, y) = \begin{cases} \dfrac{x^2 - xy}{x + y}, & \text{if } (x, y) \neq (0, 0) \\ 0, & \text{if } (x, y) = (0, 0) \end{cases}$, find $f_x(0, 0)$ and $f_y(0, 0)$.

4. If $f(x, y) = \begin{cases} \dfrac{x^3 + y^3}{x - y}, & x \neq y \\ 0, & x = y \end{cases}$, show that the function is discontinuous at the origin but possesses partial derivatives f_x and f_y at every point, including the origin.

5. If $f(x, y) = \begin{cases} xy \tan(y/x), & (x, y) \neq (0, 0) \\ 0, & (x, y) = (0, 0) \end{cases}$, show that $xf_x + yf_y = 2f$.

6. Calculate f_x, f_y, $f_x(0, 0)$, $f_y(0, 0)$ for the following:

   (*i*) $f(x, y) = \begin{cases} \dfrac{x^3 - y^3}{x^2 + y^2}, & x \neq 0, y \neq 0 \\ 0, & x = 0 = y \end{cases}$

   (*ii*) $f(x, y) = \begin{cases} \dfrac{xy}{\sqrt{x^2 + y^2}}, & \text{if } x^2 + y^2 \neq 0 \\ 0, & \text{if } x = y = 0. \end{cases}$

7. Show that the function

$$f(x, y) = \begin{cases} \dfrac{x^2 y}{x^4 + y^2}, & x^2 + y^2 \neq 0 \\ 0, & x = 0 = y \end{cases}$$

possesses first partial derivatives everywhere, including the origin, but the function is discontinuous at the origin.

8. If $f(x, y) = \sqrt{|xy|}$, find $f_x(0, 0), f_y(0, 0)$.

3.1 A Mean Value Theorem

If f_x exists throughout a neighbourhood of a point (a, b) and $f_y(a, b)$ exists then for any point $(a + h, b + k)$ of this neighbourhood,

$$f(a + h, b + k) - f(a, b) - hf_x(a + \theta h, b + k) + k[f_y(a, b) + \eta]$$

where $0 < \theta < 1$, and η is a function of k, tending to zero with k.

Now

$$f(a + h, b + k) - f(a, b) - f(a + h, b + k) - f(a, b + k) + f(a, b + k) - f(a, b) \qquad \text{...(1)}$$

CHAPTER 15

Since f_x exists in a neighbourhood of (a, b), therefore by Lagrange's mean value theorem,

$$f(a+h, b+k) - f(a, b+k) - hf_x(a+\theta h, b+k), 0<\theta<1 \qquad \text{...(2)}$$

Also $f_y(a, b)$ exists, so that

$$\lim_{k \to 0} \frac{f(a, b+k) - f(a, b)}{k} = f_y(a, b)$$

$$\Rightarrow \qquad f(a, b+k) - f(a, b) = k[f_y(a, b) + \eta] \qquad \text{...(3)}$$

where η is a function of k and tends to zero as $k \to 0$.

From equations (1), (2) and (3) we get the required result.

3.2 A Sufficient Condition for Continuity

A sufficient condition that a function f be continuous at (a, b) is that one of the partial derivatives exists and is bounded in a neighbourhood of (a, b) and that the other exists at (a, b).

Let f_x exists and be bounded in a neighbourhood of (a, b) and let $f_y(a, b)$ exists, then for any point $(a+h, b+k)$ of this neighbourhood we have (§ 3.1)

$$f(a+h, b+k) - f(a, b) - hf_x(a+\theta h, b+k) + k[f_y(a, b) + \eta]$$

where $0<\theta<1$, and $\eta \to 0$ as $k \to 0$.

Proceeding to limits as $(h, k) \to (0, 0)$, since $f_x(a+\theta h, b+k)$ is bounded, we have

$$\lim_{(h, k) \to (0, 0)} f(a+h, b+k) = f(a, b)$$

$\Rightarrow$ $\quad f$ is continuous at (a, b).

Note: A sufficient condition that a function be continuous in a *closed region* is that both the partial derivatives exist and are bounded throughout the region.

4. DIFFERENTIABILITY

Let (x, y), $(x+\delta x, y+\delta y)$ be two neighbouring points in the domain of definition of a function f. The change δf in the function as the point changes from (x, y) to $(x+\delta x, y+\delta y)$ is given by

$$\delta f = f(x+\delta x, y+\delta y) - f(x, y)$$

The function f is said to be *differentiable* at (x, y) if the change δf can be expressed in the form

$$\delta f = A\,\delta x + B\,\delta y + \delta x\,\phi(\delta x, \delta y) + \delta y\,\psi(\delta x, \delta y) \qquad \text{...(1)}$$

where A and B are constants independent of δx, δy and ϕ, ψ are functions of δx, δy tending to zero as δx, δy tend to zero simultaneously.

Also, $A\delta x + B\delta y$ is then called the *differential* of f at (x, y) and is denoted by df. Thus

$$df = A\delta x + B\delta y$$

From (1) when $(\delta x, \delta y) \to (0, 0)$, we get

$$f(x + \delta x,\ y + \delta y) - f(x,\ y) \to 0$$

or

$$f(x + \delta x,\ y + \delta y) \to f(x,\ y)$$

$\Rightarrow$ The function f is continuous at (x, y)

Thus every *differentiable function is continuous.*

Again from (1), when $\delta y = 0$ (*i.e.*, y remains constant)

$$\delta f = A\ \delta x + \delta x\ \phi(\delta x,\ 0)$$

Dividing by δx and proceeding to limits as $\delta x \to 0$, we get

$$\frac{\partial f}{\partial x} = A$$

Similarly,

$$\frac{\partial f}{\partial y} = B$$

Thus, the constants A and B are respectively the partial derivatives of f with respect to x and y.

Hence, *a function which is differentiable at a point possesses the first order partial derivatives there at.*

Converse, of course is not true, so that functions exist which are continuous and may even possess partial derivatives at a point but are not differentiable there at (see example 12),

Again the differential of f is given by

$$df = A\,\delta x + B\,\delta y = \frac{\partial f}{\partial x}\delta x + \frac{\partial f}{\partial y}\delta y$$

Taking $f = x$, we get $dx = \delta x$.

Similarly taking $f = y$, we obtain $dy = \delta y$.

Thus, the differentials dx, dy of x, y are respectively δx and δy, and

$$df = \frac{\partial f}{\partial x}\,dx + \frac{\partial f}{\partial y}\,dy = f_x\,dx + f_y\,dy \qquad ...(2)$$

is the differential of f at (x, y).

Notes:

1. If we replace δx, δy, by h, k in equation (1) we say that the function is differentiable at a point (a, b) of the domain of definition if df can be expressed as

$$\begin{aligned} df &= f(a + h, b + k) - f(a, b) \\ &= Ah + Bk + h\phi(h, k) + k\psi(h, k) \qquad ...(3) \end{aligned}$$

where $A = f_x$, $B = f_y$ and ϕ, ψ are function of h, k tending to zero as h, k tend to zero simultaneously.

2. We have seen that a function differentiable at a point is necessarily continuous and possesses partial derivatives there at. Not only that, we talk of differentiability at a point of a function only when it is continuous and has partial derivatives there at, for it is only then that it can be expressed in the form of equation (1).

Let a function f and its partial derivatives f_x, f_y be continuous at a point (x, y) of its domain of definition, and let

$$\delta f = f(x + \delta x, y + \delta y) - f(x, y)$$
$$= \{f(x + \delta x, y + \delta y) - f(x, y + \delta y)\} + \{f(x, y + \delta y) - f(x, y)\}$$

Using Lagrange's mean value theorem of one variable, we get

$$\delta f = \delta x\, f_x(x + 0_1\, \delta x, y + \delta y) + \delta y\, f_y(x, y + \theta_2\, \delta y)$$

where $0 < \theta_1 < 1$, $0 < \theta_2 < 1$.

Since f_x, f_y are continuous at (x, y) therefore when $(\delta x, \delta y) \to (0, 0)$, we get

$$\delta f = (f_x + \phi)\, \delta x + (f_y + \psi)\, \delta y$$

when ϕ and ψ tend to zero as $(\delta x, \delta y) \to (0, 0)$

$$\therefore \qquad \delta f = f_x\, \delta x + f_y\, \delta y + \delta x \phi + \delta y \psi$$

We now give an example to show that a function may be continuous and possess partial derivatives at a point and still may not be differentiable there at.

Example 12. Let

$$f(x, y) = \begin{cases} \dfrac{x^3 - y^3}{x^2 + y^2}, & (x, y) \neq (0, 0) \\ 0, & (x, y) = (0, 0) \end{cases}$$

Put $x = r\cos\theta$, $y = r\sin\theta$.

$$\therefore \qquad \left| \frac{x^3 - y^3}{x^2 + y^2} \right| = \left| r(\cos^3\theta - \sin^3\theta) \right| \leq 2|r| = 2\sqrt{x^2 + y^2} < \varepsilon,$$

if

$$x^2 < \frac{\varepsilon^2}{8}, \quad y^2 < \frac{\varepsilon^2}{8}$$

or, if

$$|x| < \frac{\varepsilon}{2\sqrt{2}}, \quad |y| < \frac{\varepsilon}{2\sqrt{2}}$$

$$\therefore \qquad \left| \frac{x^3 - y^3}{x^2 + y^2} - 0 \right| < \varepsilon, \text{ when } |x| < \frac{\varepsilon}{2\sqrt{2}}, |y| < \frac{\varepsilon}{2\sqrt{2}}$$

$$\Rightarrow \qquad \lim_{(x, y) \to (0, 0)} \frac{x^3 - y^3}{x^2 + y^2} = 0$$

$$\Rightarrow \qquad \lim_{(x, y) \to (0, 0)} f(x, y) = f(0, 0)$$

Hence the function is continuous at $(0, 0)$.

Again,

$$f_x(0, 0) = \lim_{h \to 0} \frac{f(h, 0) - f(0, 0)}{h} = \lim_{h \to 0} \frac{h - 0}{h} = 1$$

$$f_y(0, 0) = \lim_{k \to 0} \frac{f(0, k) - f(0, 0)}{k} = \lim_{k \to 0} \frac{-k}{k} = -1$$

Thus, the function possesses partial derivatives at (0, 0).

If the function is differentiable at (0, 0), then by definition

$$df = f(h, k) - f(0, 0) = Ah + Bk + h\phi + k\psi \quad \text{...(1)}$$

when A and B are constants $(A = f_x(0, 0) = 1,\ B = f_y(0, 0) = -1)$ and ϕ, ψ tend to zero as $(h, k) \to (0, 0)$.

Putting $h = \rho \cos\theta$, $k = \rho \sin\theta$, and dividing by ρ, we get

$$\cos^3\theta - \sin^3\theta = \cos\theta - \sin\theta + \phi \cos\theta + \psi \sin\theta \quad \text{...(2)}$$

For arbitrary $\theta = \tan^{-1}(h/k)$, $\rho \to 0$ implies that $(h, k) \to (0, 0)$. Thus we get the limit,

$$\cos^3\theta - \sin^3\theta = \cos\theta - \sin\theta$$

or

$$\cos\theta \sin\theta (\cos\theta - \sin\theta) = 0$$

which is plainly impossible for arbitrary θ.

Thus, the function is not differentiable at the origin.

Note: The method used to show that the function is not differentiable, can also be used to show that the function is not continuous at (0, 0); for example,

The function f, where

$$f(x, y) = \begin{cases} \dfrac{xy}{x^2 + y^2}, & x^2 + y^2 \neq 0 \\ 0, & \text{if } x = y = 0 \end{cases}$$

is not differentiable at the origin because it is discontinuous there at.

Example 13. Show that the function f, where

$$f(x, y) = \begin{cases} \dfrac{xy}{\sqrt{x^2 + y^2}}, & \text{if } x^2 + y^2 \neq 0 \\ 0, & \text{if } x = y = 0 \end{cases}$$

is continuous, possesses partial derivative but is not differentiable at the origin.

■ As was shown in Example 9, f is continuous at the origin. Also it may be easily shown that

$$f_x(0, 0) = 0 = f_y(0, 0)$$

If the function is differentiable at the origin, then by definition

$$df = f(h, k) - f(0, 0) = Ah + Bk + h\phi + k\psi \quad \text{...(1)}$$

where $A = f_x(0, 0) = 0$, $B = f_y(0, 0) = 0$, and ϕ, ψ tend to zero as $(h, k) \to (0, 0)$.

$$\therefore \qquad \frac{hk}{\sqrt{h^2 + k^2}} = h\phi + k\psi \qquad \ldots(2)$$

Putting $k = mh$ and letting $h \to 0$, we get

$$\frac{m}{\sqrt{1 + m^2}} = \lim_{h \to 0} (\phi + m\psi) = 0$$

which is impossible for arbitrary m.

Hence, the function is not differentiable at (0, 0).

Note: If we put $h = r\cos\theta, k = r\sin\theta$ in (2) we get

$$\cos\theta \sin\theta = \phi\cos\theta + \psi\sin\theta$$

For arbitrary θ, $r \to 0$ implies $(h, k) \to (0, 0)$.

Thus when $r \to 0$, we get

$$\cos\theta \cdot \sin\theta = 0$$

which is impossible for arbitrary θ. So f is not differentiable at the origin.

Ex. 1. Show that the function f, where

$$f(x, y) = \begin{cases} x\sin 1/x + y\sin 1/y, & xy \neq 0 \\ x\sin 1/x, & y = 0, x \neq 0 \\ y\sin 1/y, & x = 0, y \neq 0 \\ 0, & x = 0 = y \end{cases}$$

is continuous but not differentiable at the origin.

Ex. 2. Show that the function $|x| + |y|$ is continuous, but not differentiable at the origin.

Ex. 3. Discuss the following functions for continuity and differentiability at the origin.

(*i*) $f(x, y) = \dfrac{xy^2}{x^2 + y^2}$ when $(x, y) \neq (0, 0)$ and $f(0, 0) = 0$

(*ii*) $f(x, y) = y\sin 1/x$, if $x \neq 0$, $f(0, y) = y$.

4.1 A Sufficient Condition for Differentiability

Theorem 2. *If (a, b) be a point of the domain of definition of a function f such that*

(*i*) *f_x is continuous at (a, b),*

(*ii*) *f_y exists at (a, b),*

then f is differentiable at (a, b).

The condition (*i*) implies that f_x exists in a certain neighbourhood $(a - \delta, a + \delta; b - \delta, b + \delta)$ of (a, b). Let $(a + h, b + k)$ be a point of this neighbourhood. Thus

$$df = f(a+h, b+k) - f(a, b)$$
$$= f(a+h, b+k) - f(a, b+k) + f(a, b+k) - f(a, b) \qquad ...(1)$$

Since f_x exists in $(a-\delta, a+\delta; b-\delta, b+\delta)$, applying Lagrange's mean value theorem, we get

$$f(a+h, b+k) - f(a, b+k) = hf_x(a+\theta h, b+k) \qquad ...(2)$$

where $0 < \theta < 1$, and depends on h and k.

Again, since f_x is continuous at (a, b), therefore

$$\lim_{(h,k)\to(0,0)} f_x(a+\theta h, b+k) = f_x(a, b)$$

so that we can write

$$f_x(a+\theta h, b+k) = f_x(a, b) + \phi(h, k) \qquad ...(3)$$

where $\phi(h, k) \to 0$ as $(h, k) \to (0, 0)$.

Again, since by condition (*ii*), $f_y(a, b)$ exists, therefore

$$\lim_{k\to 0} \frac{f(a, b+k) - f(a, b)}{k} = f_y(a, b)$$

so that we can write

$$\frac{f(a, b+k) - f(a, b)}{k} = f_y(a, b) + \psi(k) \qquad ...(4)$$

where $\psi(k) \to 0$ as $k \to 0$.

$\therefore$ From (1), (2), (3) and (4), we get

$$df = hf_x(a, b) + kf_y(a, b) + h\phi(h, k) + k\psi(k)$$

$\Rightarrow$ f is differentiable at (a, b).

Note: In a similar way it can be shown that f is differentiable at (a, b), if f_x exists and f_y is continuous at (a, b). In fact, one of the partial derivatives is to be continuous and the other merely to exist at the point.

Remark: We have shown that the condition of existence of one partial derivative and the continuity of the other is sufficient to ensure that the function is differentiable but with the help of an example (Example I below) we now show that the condition of continuity is not necessary so that function may be differentiable even though none of the partial derivatives is continuous. However, if the function is not differentiable at a point, the partial derivatives cannot be continuous there at (Example II).

Example I. Consider the function

$$f(x, y) = \begin{cases} x^2 \sin 1/x + y^2 \sin 1/y, & \text{if } xy \neq 0 \\ x^2 \sin 1/x, & \text{if } x \neq 0 \text{ and } y = 0 \\ y^2 \sin 1/y, & \text{if } x = 0 \text{ and } y \neq 0 \\ 0, & \text{if } x = y = 0 \end{cases}$$

■ The partial derivatives,

$$f_x(x, y) = \begin{cases} 2x \sin 1/x - \cos 1/x, & \text{if } x \neq 0 \\ 0, & \text{if } x = 0 \end{cases}$$

$$f_y(x, y) = \begin{cases} 2y \sin 1/y - \cos 1/y, & \text{if } y \neq 0 \\ 0, & \text{if } y = 0 \end{cases}$$

are discontinuous at the origin, so that both the partial derivatives exist at the origin, but none is continuous there at.

Let us show that the function is differentiable at the origin. Here,

$$\begin{aligned} f(h, k) - f(0, 0) &= h^2 \sin 1/h + k^2 \sin 1/k \\ &= 0h + 0k + h(h \sin 1/h) + k(k \sin 1/k) \end{aligned}$$

Now ($h \sin 1/h$) and ($k \sin 1/k$) both tend to zero when $(h, k) \to (0, 0)$ so that f is differentiable at the origin.

Example II. Prove that the function

$$f(x, y) = \sqrt{|xy|}$$

is not differentiable at the point (0, 0), but that f_x and f_y both exist at the origin and have the value 0. Hence deduce that these two partial derivatives are continuous except at the origin.

■ Now at (0, 0),

$$f_x(0, 0) = \lim_{h \to 0} \frac{f(h, 0) - f(0, 0)}{h} = \lim_{h \to 0} \frac{0}{h} = 0$$

$$f_y(0, 0) = \lim_{k \to 0} \frac{f(0, k) - f(0, 0)}{k} = \lim_{k \to 0} \frac{0}{k} = 0$$

If the function is differentiable at (0, 0) then by definition

$$f(h, k) - f(0, 0) = 0h + 0k + h\phi + k\psi$$

where ϕ and ψ are functions of h and k, and tend to zero as $(h, k) \to (0, 0)$.

Putting $h = \rho \cos\theta$, $k = \rho \sin\theta$ and dividing by ρ, we get

$$|\cos\theta \sin\theta|^{1/2} = \phi \cos\theta + \psi \sin\theta$$

Now for arbitrary θ, $\rho \to 0$ implies that $(h, k) \to (0, 0)$.

Taking the limit as $\rho \to 0$, we get

$$|\cos\theta \sin\theta|^{1/2} = 0,$$

which is impossible for all arbitrary θ.

Hence, the function is not differentiable at (0, 0) and consequently the partial derivatives f_x, f_y cannot be continuous at (0, 0), for otherwise the function would be differentiable there at,

Let us now see that it is actually so

For $(x, y) \neq (0, 0)$.

$$f_x(x, y) = \lim_{h \to 0} \frac{f(x+h, y) - f(x, y)}{h}$$

$$= \lim_{h \to 0} \frac{\sqrt{|x+h||y|} - \sqrt{|x||y|}}{h}$$

$$= \lim_{h \to 0} \sqrt{|y|} \frac{|x+h| - |x|}{h\left[\sqrt{|x+h|} + \sqrt{|x|}\right]}$$

Now as $h \to 0$, we can take $x + h > 0$, *i.e.*, $|x + h| = x + h$, when $x > 0$ and $x + h < 0$ or $|x + h| = -(x + h)$, when $x < 0$.

$$\therefore \qquad f_x(x, y) = \begin{cases} \dfrac{1}{2}\sqrt{\dfrac{|y|}{|x|}}, & \text{when } x > 0 \\ -\dfrac{1}{2}\sqrt{\dfrac{|y|}{|x|}}, & \text{when } x < 0 \end{cases}$$

Similarly,

$$f_y(x, y) = \begin{cases} \dfrac{1}{2}\sqrt{\dfrac{|x|}{|y|}}, & \text{when } y > 0 \\ -\dfrac{1}{2}\sqrt{\dfrac{|x|}{|y|}}, & \text{when } y < 0 \end{cases}$$

which are, obviously, not continuous at the origin.

Example 14. Show that the function f, where

$$f(x, y) = \begin{cases} xy\dfrac{x^2 - y^2}{x^2 + y^2}, & \text{if } x^2 + y^2 \neq 0 \\ 0, & \text{if } x = y = 0 \end{cases}$$

is differentiable at the origin.

■ It may be easily shown that

$$f_x(0, 0) = 0 = f_y(0, 0)$$

Also when $x^2 + y^2 \neq 0$,

$$|f_x| = \frac{|x^4 y + 4x^2 y^3 - y^5|}{(x^2 + y^2)^2} \leq \frac{6(x^2 + y^2)^{5/2}}{(x^2 + y^2)^2} = 6(x^2 + y^2)^{1/2}$$

Evidently

$$\lim_{(x,y)\to(0,0)} f_x(x, y) = 0 = f_x(0, 0)$$

Thus f_x is continuous at (0, 0) and $f_y(0, 0)$ exists.

$\Rightarrow$ f is differentiable at (0, 0).

4.2 Algebra of Differentiable Functions

If f and g are two functions differentiable at (a, b), then $f \pm g$, fg are differentiable at (a, b); f/g is differentiable at (a, b), if $g(a, b) \neq 0$, and

$$d(f \pm g) = df \pm dg$$

$$d(fg) = gdf + fdg$$

$$d(f/g) = (gdf + fdg)/g^2.$$

5. PARTIAL DERIVATIVES OF HIGHER ORDER

If a function f has partial derivatives of the first order at each point (x, y) of a certain region, then f_x, f_y are themselves functions of x, y and may also possess partial derivatives. These are called *second order partial derivatives of* f and are denoted by

$$\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial^2 f}{\partial x^2} = f_{xx} = f_{x^2}$$

$$\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial^2 f}{\partial y^2} = f_{yy} = f_{y^2}$$

$$\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial^2 f}{\partial x \partial y} = f_{xy}$$

$$\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial^2 f}{\partial y \partial x} = f_{yx}$$

In a similar manner higher order partial derivatives are defined. For example $\dfrac{\partial^3 f}{\partial x \partial x \partial y} = f_{xxy}$ and so on.

The second order partial derivatives at a particular point (a, b) are often denoted by

$$\left[\frac{\partial^2 f}{\partial x^2}\right]_{(a,\,b)}, \frac{\partial^2 f(a, b)}{\partial x^2}, f_{xx}(a, b) \text{ or } f_{x^2}(a, b)$$

$$\left[\frac{\partial^2 f}{\delta x \partial y}\right]_{(a,\,b)}, \frac{\partial^2 f(a, b)}{\partial x \partial y} \text{ or } f_{xy}(a, b)$$

and so on.

Thus

$$f_{xx}(a, b) = \lim_{h\to 0} \frac{f_x(a+h, b) - f_x(a, b)}{h}$$

$$f_{xy}(a, b) = \lim_{h\to 0} \frac{f_y(a+h, b) - f_y(a, b)}{h}$$

$$f_{yx}(a, b) = \lim_{k\to 0} \frac{f_x(a, b+k) - f_x(a, b)}{k}$$

$$f_{yy}(a, b) = \lim_{k\to 0} \frac{f_y(a, b+k) - f_y(a, b)}{k}$$

in case the limits exist.

5.1 Change in the Order of Partial Derivation

In most of the cases that occur in practice, a partial derivative has the same value in whatever order the different operations are performed. Thus, for example, it is usually found that

$$f_{xy} = f_{yx}, \quad f_{xyx} = f_{xxy}, \quad f_{xyxy} = f_{xxyy}$$

and one is often tempted to believe that it is always so. But it is not the case and there is no *a priori* reason why they should be equal. Let us now see why f_{xy} may be different from f_{yx} at some point (a, b) of the region.

Now

$$f_{xy}(a, b) = \lim_{h\to 0} \frac{f_y(a+h, b) - f_y(a, b)}{h}$$

$$= \lim_{h\to 0} \frac{1}{h}\left[\lim_{k\to 0} \frac{f(a+h, b+k) - f(a+h, b)}{k} - \lim_{k\to 0} \frac{f(a, b+k) - f(a, b)}{k}\right]$$

$$= \lim_{h\to 0} \lim_{k\to 0} \frac{f(a+h, b+k) - f(a+h, b) - f(a, b+k) + f(a, b)}{hk} = \lim_{h\to 0} \lim_{k\to 0} \frac{\phi(h, k)}{hk}$$

where $\phi(h, k) = f(a+h, b+k) - f(a+h, b) - f(a, b+k) + f(a, b)$.

Similarly,

$$f_{yx}(a, b) = \lim_{k\to 0} \lim_{h\to 0} \frac{\phi(h, k)}{hk}$$

Thus we see that $f_{xy}(a, b)$ and $f_{yx}(a, b)$ are the repeated limits of the same expression taken in different orders. There is therefore no *a priori* reason why they should always be equal.

Let us consider an example to show that f_{xy} may be different from f_{yx}.

Example 15. Let

$$f(x, y) = \frac{xy(x^2 - y^2)}{x^2 + y^2}, \ (x, y) \neq (0, 0), \ f(0, 0) = 0, \text{ then}$$

show that at the origin $f_{xy} \neq f_{yx}$.

■ Now

$$f_{xy}(0, 0) = \lim_{h\to 0} \frac{f_y(h, 0) - f_y(0, 0)}{h}$$

$$f_y(0, 0) = \lim_{k\to 0} \frac{f(0, k) - f(0, 0)}{k} = \lim_{k\to 0} \frac{0}{k} = 0$$

$$f_y(h, 0) = \lim_{k\to 0} \frac{f(h, k) - f(h, 0)}{k} = \lim_{k\to 0} \frac{hk(h^2 - k^2)}{k \cdot (h^2 + k^2)} = h$$

$\therefore$
$$f_{xy} = \lim_{h\to 0} \frac{h - 0}{h} = 1$$

Again

$$f_{yx}(0, 0) = \lim_{k\to 0} \frac{f_x(0, k) - f_x(0, 0)}{k}$$

But

$$f_x(0, 0) = \lim_{h\to 0} \frac{f(h, 0) - f(0, 0)}{h} = 0$$

$$f_x(0, k) = \lim_{h\to 0} \frac{f(h, k) - f(0, k)}{h} = \lim_{h\to 0} \frac{hk(h^2 - k^2)}{h(h^2 + k^2)} = -k$$

$\therefore$
$$f_{yx}(0, 0) = \lim_{k\to 0} \frac{-k - 0}{k} = -1$$

$\therefore$
$$f_{xy}(0, 0) \neq f_{yx}(0, 0).$$

Example 16. Examine the equality of f_{xy} and f_{yx}, where

$$f(x, y) = x^3 y + e^{xy^2}$$

■ Now

$$f_y = x^3 + 2xye^{xy^2}$$

$\therefore$
$$f_{xy} = 3x^2 + 2ye^{xy^2} + 2xy^3e^{xy^2}$$

Again

$$f_x = 3x^2y + y^2e^{xy^2}$$

$$f_{yx} = 3x^2 + 2ye^{xy^2} + 2xy^3e^{xy^2}$$

$\Rightarrow$
$$f_{xy} = f_{yx}.$$

EXERCISE

1. Verify that $f_{xy} = f_{yx}$ for the functions:

   (*a*) $\dfrac{2x - y}{x + y}$, (*b*) $x \tan xy$, (*c*) $\cosh (y + \cos x)$, (*d*) x^y

   indicating possible exceptional points and investigate these points.

2. Show that $z = \log \{(x - a)^2 + (y - b)^2\}$ satisfies $\dfrac{\partial^2 z}{\partial x^2} + \dfrac{\partial^2 z}{\partial y^2} = 0$, except at (a, b).

3. Show that $z = x \cos (y/x) + \tan (y/x)$ satisfies $x^2 z_{xx} + 2xyz_{xy} + y^2 z_{yy} = 0$, except at points for which $x = 0$.

4. Prove that $f_{xy} \neq f_{yx}$ at the origin for the function:

$$f(x, y) = x^2 \tan^{-1}(y/x) - y^2 \tan^{-1}(x/y),\ x \neq 0,\ y \neq 0$$
$$f(x, y) = 0, \text{ elsewhere.}$$

5. If $f(x, y, z) = (x^2 + y^2 + z^2)^{-1/2}$, show that

$$\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2} = 0.$$

6. Examine for the change in the order of derivation at the origin for the functions:

   (*i*) $f(x, y) = e^x (\cos y + x \sin y)$

   (*ii*) $f(x, y) = \sqrt{x^2 + y^2} \sin 2\phi$,

   where $f(0, 0) = 0$ and $\phi = \tan^{-1}(y/x)$,

   (*iii*) $f(x, y) = \left| x^2 - y^2 \right|$.

7. Examine the equality of $f_{xy}(0, 0)$ and $f_{yx}(0, 0)$ for the function:

$$f(x, y) = (x^2 + y^2) \tan^{-1}(y/x),\quad x \neq 0,\ f(0, y) = \pi y^2/2.$$

8. Given $u = e^x \cos y + e^y \sin z$, find all first partial derivatives and verify that

$$\frac{\partial^2 u}{\partial x \partial y} = \frac{\partial^2 u}{\partial y \partial x};\ \frac{\partial^2 u}{\partial x \partial z} = \frac{\partial^2 u}{\partial z \partial x};\ \frac{\partial^2 u}{\partial y \partial z} = \frac{\partial^2 u}{\partial z \partial y}.$$

5.2 Sufficient Conditions for the Equality of f_{xy} and f_{yx}

As was said earlier there is no *a priori* reason why f_{xy} and f_{yx} should always be equal. We now give two theorems the object of which is to set out precisely under what conditions it is safe to assume that $f_{xy} = f_{yx}$ at a point, *i.e.*, **sufficient conditions** for the equality of f_{xy} and f_{yx}.

Theorem 3. ***Young's theorem.*** *If f_x and f_y are both differentiable at a point (a, b) of the domain of definition of a function f, then*

$$f_{xy}(a, b) = f_{yx}(a, b)$$

The differentiability of f_x and f_y at (a, b) implies that they exist in a certain neighbourhood of (a, b) and that all the second order partial derivatives $f_{xx}, f_{xy}, f_{yx}, f_{yy}$ exist at (a, b).

We prove the theorem by taking equal increment h both for x and y and calculating $\phi(h, h)$ in two different ways.

Let $(a + h, b + h)$ be a point of this neighbourhood. Consider

$$\phi(h, h) = f(a + h, b + h) - f(a + h, b) - f(a, b + h) + f(a, b)$$

$$G(x) = f(x, b + h) - f(x, b)$$

so that

$$\phi(h, h) = G(a + h) - G(a) \qquad ...(1)$$

Since f_x exists in a neighbourhood of (a, b), the function $G(x)$ is derivable in $]a, a + h[$ and therefore by Lagrange's mean value theorem, we get from (1),

$$\phi(h, h) = hG'(a + \theta h), \quad 0 < \theta < 1$$

$$= h\{f_x(a + \theta h, b + h) - f_x(a + \theta h, b)\} \qquad ...(2)$$

Again, since f_x is differentiable at (a, b), we have

$$f_x(a + \theta h, b + h) - f_x(a, b) = \theta h f_{xx}(a, b) + h f_{yx}(a, b) + \theta h \phi_1(h, h) + h\psi_1(h, h) \qquad ...(3)$$

and

$$f_x(a + \theta h, b) - f_x(a, b) = \theta h f_{xx}(a, b) + \theta h \phi_2(h, h) \qquad ...(4)$$

where ϕ_1, ψ_1, ϕ_2 all tend to zero as $h \to 0$.

From equations (2), (3), and (4), we get

$$\phi(h, h)/h^2 = f_{yx}(a, b) + \theta\phi_1(h, h) + \psi_1(h, h) - \theta\phi_2(h, h) \qquad ...(5)$$

By a similar argument, on considering

$$H(y) = f(a + h, y) - f(a, y)$$

we can show that

$$\phi(h, h)/h^2 = f_{xy}(a, b) + \phi_3(h, h) + \theta'\psi_2(h, h) - \theta'\psi_3(h, h) \qquad ...(6)$$

where ϕ_3, ψ_2, ψ_3 all tend to zero as $h \to 0$.

On taking the limit as $h \to 0$, we obtain from equations (5) and (6)

$$\lim_{h \to 0} \frac{\phi(h, h)}{h^2} = f_{xy}(a, b) = f_{yx}(a, b).$$

Note: An alternative set of conditions which involves only the existence of one of the second order partial derivatives of f at (a, b) provided we assume also its continuity, is made in the following next theorem.

Theorem 4. ***Schwarz's theorem.*** *If f_y exists in a certain neighbourhood of a point (a, b) of the domain of definition of a function f, and f_{yx} is continuous at (a, b), then $f_{xy}(a, b)$ exists and is equal to $f_{yx}(a, b)$.*

Under the given conditions, f_x, f_y, and f_{yx} all exist in a certain neighbourhood of (a, b). Let $(a + h, b + k)$ be a point of this neighbourhood.

Consider

$$\phi(h, k) = f(a+h, b+k) - f(a+h, b) - f(a, b+k) + f(a, b)$$
$$G(x) = f(x, b+k) - f(x, b)$$

so that

$$\phi(h, k) = G(a+h) - G(a) \qquad ...(1)$$

Since f_x exists in a neighbourhood of (a, b), the function $G(x)$ is derivable in $]a, a+h[$, and therefore by Lagrange's mean value theorem, we get from (1)

$$\phi(h, k) = hG'(a+\theta h), \quad 0 < \theta < 1$$
$$= h\{f_x(a+\theta h, b+k) - f_x(a+\theta h, b)\} \qquad ...(2)$$

Again, since f_{yx} exists in a neighbourhood of (a, b), the function f_x is derivable with respect to y in $]b, b+k[$, and therefore by Lagrange's mean value theorem, we get from (2)

$$\phi(h, k) = hkf_{yx}(a+\theta h, b+\theta' k), \quad 0 < \theta' < 1$$

or

$$\frac{1}{h}\left\{\frac{f(a+h, b+k) - f(a+h, b)}{k} - \frac{f(a, b+k) - f(a, b)}{k}\right\} = f_{yx}(a+\theta h, b+\theta' k)$$

Proceeding to limits when $k \to 0$, since f_y and f_{yx} exist in a neighbourhood of (a, b), we get

$$\frac{f_y(a+h, b) - f_y(a, b)}{h} = \lim_{k\to 0} f_{yx}(a+\theta h, b+\theta' k)$$

Again, taking limits as $h \to 0$, since f_{yx} is continuous at (a, b), we get

$$f_{xy}(a, b) = \lim_{h\to 0}\lim_{k\to 0} f_{yx}(a+\theta h, b+\theta' k) = f_{yx}(a, b)$$

Notes:

1. If f_{xy} *and* f_{yx} *are both continuous at* (a, b), *then* $f_{xy}(a, b) = f_{yx}(a, b)$, for the assumption of continuity of both these derivatives is a wider assumption than those required for proving either Theorem 3 or Theorem 4.
2. If the conditions of Young's or Schwarz's theorem are satisfied then $f_{xy} = f_{yx}$ at a point (a, b). But if the conditions are not satisfied, we cannot draw any conclusion regarding the equality of f_{xy} and f_{yx} they may or may not be equal (see examples 17 and 18). Thus the conditions are *sufficient* but *not necessary*.

Example 17. Show that for the function

$$f(x, y) = \begin{cases} \dfrac{x^2y^2}{x^2+y^2}, & (x, y) \neq (0, 0) \\ 0, & (x, y) = (0, 0) \end{cases}$$

$f_{xy}(0, 0) = f_{yx}(0, 0)$, even though the conditions of Schwarz's theorem and also of Young's theorem are not satisfied.

■ Now

$$f_x(0, 0) = \lim_{x\to 0} \frac{f(x, 0) - f(0, 0)}{x} = 0$$

Similarly, $f_y(0, 0) = 0$.

Also, for $(x, y) \neq (0, 0)$,

$$f_x(x, y) = \frac{(x^2 + y^2) \cdot 2xy^2 - x^2y^2 \cdot 2x}{(x^2 + y^2)^2} = \frac{2xy^4}{(x^2 + y^2)^2}$$

$$f_y(x, y) = \frac{2x^4 y}{(x^2 + y^2)^2}$$

Again

$$f_{yx}(0, 0) = \lim_{y \to 0} \frac{f_x(0, y) - f_x(0, 0)}{y} = 0$$

and

$f_{xy}(0, 0) = 0$, so that $f_{xy}(0, 0) = f_{yx}(0, 0)$

For $(x, y) \neq (0, 0)$, we have

$$f_{yx}(x, y) = \frac{8xy^3(x^2 + y^2)^2 - 2xy^4 \cdot 4y(x^2 + y^2)}{(x^2 + y^2)^4} = \frac{8x^3y^3}{(x^2 + y^2)^3}$$

and it may be easily shown (by putting $y = mx$) that

$$\lim_{(x, y) \to (0, 0)} f_{yx}(x, y) \neq 0 = f_{yx}(0, 0)$$

so that f_{yx} is not continuous at (0, 0), *i.e.*, the conditions of Schwarz's theorem are not satisfied.

Let us now show that the conditions of Young's theorem are also not satisfied.

Now

$$f_{xx}(0, 0) = \lim_{x \to 0} \frac{f_x(x, 0) - f_x(0, 0)}{x} = 0$$

Also f_x is differentiable at (0, 0) if

$$f_x(h, k) - f_x(0, 0) = f_{xx}(0, 0) \cdot h + f_{yx}(0, 0) \cdot k + h\phi + k\psi$$

or

$$\frac{2hk^4}{(h^2 + k^2)^2} = h\phi + k\psi$$

where ϕ, ψ tend to zero as $(h, k) \to (0, 0)$.

Putting $h = \rho \cos\theta$ and $k = \rho \sin\theta$, and dividing by ρ, we get

$$2 \cos\theta \sin^4\theta = \cos\theta \cdot \phi + \sin\theta\psi$$

and $(h, k) \to (0, 0)$ is same thing as $\rho \to 0$ and θ is arbitrary. Thus proceeding to limits, we get

$$2 \cos\theta \sin^4\theta = 0$$

which is impossible for arbitrary θ.

$\Rightarrow \quad f_x$ is not differentiable at (0, 0)

Similarly, it may be shown that f_y is not differentiable at (0, 0).

Thus the conditions of Young's theorem are also not satisfied but, as shown above,

$$f_{xy}(0, 0) = f_{yx}(0, 0).$$

Example 18. Show that the function

$$f(x, y) = \frac{xy(x^2 - y^2)}{x^2 + y^2}, \quad (x, y) \neq (0, 0)$$

$$f(0, 0) = 0$$

does not satisfy the conditions of Schwarz's theorem and

$$f_{xy}(0, 0) \neq f_{yx}(0, 0)$$

■ It may be shown, as in example 15, that

$$f_{xy}(0, 0) = 1, \; f_{yx}(0, 0) = -1$$

so that

$$f_{xy}(0, 0) \neq f_{yx}(0, 0)$$

Now, for $(x, y) \neq (0, 0)$ we have

$$f_x(x, y) = \frac{(x^2 + y^2)y(3x^2 - y^2) - 2x^2y(x^2 - y^2)}{(x^2 + y^2)^2} = \frac{y\{x^4 + 4x^2y^2 - y^4\}}{(x^2 + y^2)^2}$$

$$\therefore \quad f_{yx}(x, y) = \frac{(x^2 + y^2)^2\{x^4 + 12x^2y^2 - 5y^4\} - 4y^2(x^2 + y^2)\{x^4 + 4x^2y^2 - y^4\}}{(x^2 + y^2)^4}$$

$$= \frac{x^6 + 9x^4y^2 - 9x^2y^4 - y^6}{(x^2 + y^2)^3}.$$

By putting $y = mx$ or $x = r\cos\theta,\; y = r\sin\theta,$ it may be shown that

$$\lim_{(x, y)\to(0, 0)} f_{yx}(x, y) \neq -1 = f_{yx}(0, 0).$$

Thus f_{yx} is not continuous at (0, 0).

It may similarly be shown that f_{xy} is also not continuous at (0, 0).

Thus, the conditions of Schwarz's theorem are not satisfied.

6. DIFFERENTIALS OF HIGHER ORDER

Let $z = f(x, y)$ be a function of two independent variables x and y, defined in a domain N and let it be differentiable at a point (x, y) of the domain. The first differential of z at (x, y), denoted by dz is given by

$$dz = \frac{\partial z}{\partial x}dx + \frac{\partial z}{\partial y}dy \qquad \text{...(1)}$$

If dx and dy are regarded as constants and if $\frac{\partial z}{\partial x}$ and $\frac{\partial z}{\partial y}$ are differentiable at (x, y) then dz is a function of x and y and is itself differentiable at (x, y). The differential of dz, called the *second differential* of z, is denoted by d^2z and is calculated in the same way as the first.

$$\therefore \qquad d^2 z = d(dz) = d\left(\frac{\partial z}{\partial x}\right)dx + d\left(\frac{\partial z}{\partial y}\right)dy \qquad ...(2)$$

Replacing z by $\frac{\partial z}{\partial x}$ and $\frac{\partial z}{\partial y}$ in (1), we get

$$d\left(\frac{\partial z}{\partial x}\right) = \frac{\partial^2 z}{\partial x^2}dx + \frac{\partial^2 z}{\partial y \partial x}dy$$

$$d\left(\frac{\partial z}{\partial y}\right) = \frac{\partial^2 z}{\partial x \partial y}dx + \frac{\partial^2 z}{\partial y^2}dy$$

Also by Young's theorem, since $\frac{\partial z}{\partial x}$ and $\frac{\partial z}{\partial y}$ are differentiable, we have

$$\frac{\partial^2 z}{\partial x \partial y} = \frac{\partial^2 z}{\partial y \partial x}$$

$$\therefore \qquad d^2 z = \frac{\partial^2 z}{\partial x^2}dx^2 + 2\frac{\partial^2 z}{\partial x \partial y}dx\, dy + \frac{\partial^2 z}{\partial y^2}dy^2 \qquad ...(3)$$

where, of course, $dx^2 = dx \cdot dx = (dx)^2$, $dy^2 = (dy)^2$

In abbreviated notation, it may be written as

$$d^2 z = \left(\frac{\partial}{\partial x}dx + \frac{\partial}{\partial y}dy\right)^2 z \qquad ...(4)$$

Again d^2z is differentiable at (x, y) if all the second order partial derivatives $\frac{\partial^2 z}{\partial x^2}, \frac{\partial^2 z}{\partial x \partial y}, \frac{\partial^2 z}{\partial y^2}$ are differentiable at (x, y). This condition also ensures the legitimacy of inverting the order of the partial derivatives with respect to x and with respect to y, and so

$$d^3 z = \frac{\partial^3 z}{\partial x^3}dx^3 + 3\frac{\partial^3 z}{\partial x^2 \partial y}dx^2\, dy + 3\frac{\partial^3 z}{\partial x \partial y^2}dx\, dy^2 + \frac{\partial^3 z}{\partial y^3}\partial y^3 = \left(\frac{\partial}{\partial x}dx + \frac{\partial}{\partial y}dy\right)^3 z \qquad ...(5)$$

Proceeding in the manner, we can define the successive differentials $d^4 z, d^5 z,$... Thus the differential of nth order, $d^n z$ exists if $d^{n-1}z$ is differentiable, which implies that all the partial derivatives of the $(n-1)$th order are differentiable. This condition also ensures the legitimacy of inverting the order of the partial derivatives with respect to x and with respect to y in the partial derivatives of order n. Thus it may be shown by Mathematical induction that

$$d^n z = \frac{\partial^n z}{\partial x^n} dx^n + n \frac{\partial^n z}{\partial x^{n-1} \partial y} dx^{n-1} dy + \frac{n(n-1)}{2!} \frac{\partial^n z}{\partial x^{n-2} \partial y^2} dx^{n-2} dy^2 + \dots + \frac{\partial^n z}{\partial y^n} dy^n$$

$$= \left(\frac{\partial}{\partial x} dx + \frac{\partial}{\partial y} dy \right)^n z.$$

Note: In the above discussion, x and y are *Independent Variables* and so dx and dy may be treated as constants. The reason for this being so is that the differentials of independent variables are the arbitrary increments of these variables, $dx = \delta x$, $dy = \delta y$.

7. FUNCTIONS OF FUNCTIONS

So far we have considered functions of the form

$$z = f(x, y, \dots)$$

where the variables $x, y, \dots$ are the independent variables. We now consider functions

$$z = f(x, y, \dots)$$

where $x, y, \dots$ are not independent variables, but are themselves functions of other independent variables $u, v, \dots$, so that

$$x = g(u, v, \dots) \text{ and } y = h(u, v, \dots)$$

To fix the ideas, we consider only two variables x and y as functions of two independent variables u and v. The method of proof is, however, general.

Theorem 5. *If $z = f(x, y)$ is a differentiable function of x, y and $x = g(u, v)$, $y = h(u, v)$ are themselves differentiable functions of the independent variables u, v, then z is a differentiable function of u, v and*

$$dz = \frac{\partial z}{\partial x} dx + \frac{\partial z}{\partial y} dy$$

just as though x, y were the independent variables.

Let (u, v), $(u + \delta u, v + \delta v)$ be two neighbouring points of the domain of definition of x and y, and (x, y), $(x + \delta x, y + \delta y)$ the corresponding points of the domain of definition of z, so that

$$\delta x = g(u + \delta u, v + \delta v) - g(u, v)$$

$$\delta y = h(u + \delta u, v + \delta v) - h(u, v)$$

The differentiability, and hence the continuity of g and h imply that

$$\delta x \to 0, \delta y \to 0, \text{ as } (\delta u, \delta v) \to (0, 0)$$

Again, since g and h are differentiable function of u and v,

$$\delta x = g_u \delta u + g_v \delta v + \phi_1 \delta u + \psi_1 \delta v \quad \text{...(1)}$$

$$\delta y = h_u \delta u + h_v \delta v + \phi_2 \delta u + \psi_2 \delta v, \quad \text{...(2)}$$

CHAPTER 15

where $\phi_1, \phi_2, \psi_1, \psi_2$ are functions of $\delta u, \delta v$, and tend to zero as,

$$(\delta u, \delta v) \to (0, 0).$$

Also, $dx = g_u\, du + g_v\, dv$, $dy = h_u\, du + h_v\, dv$.

Also, since f is a differentiable function of x, y, we have

$$\delta z = f_x\, \delta x + f_y\, \delta y + \phi_3\, \delta x + \psi_3\, \delta y, \qquad \text{...(3)}$$

where ϕ_3, ψ_3 are functions of $\delta x, \delta y$, and tend to zero as $(\delta x, \delta y) \to (0, 0)$.

From equations (1), (2), and (3) we get

$$\delta z = (f_x g_u + f_y h_u)\, \delta u + (f_x g_v + f_y h_v)\, \delta v + F_1\, \delta u + F_2\, \delta v$$

where

$$F_1 = f_x\phi_1 + f_y\phi_2 + \phi_3 g_u + \phi_3\phi_1 + \psi_3 h_u + \psi_3\phi_2$$

$$F_2 = f_x\psi_1 + f_y\psi_2 + \phi_3 g_v + \phi_3\psi_1 + \psi_3 h_v + \psi_3\psi_2$$

Since the coefficients F_1 and F_2 of $\delta u, \delta v$ tend to zero as $(\delta u, \delta v) \to (0, 0)$, therefore z is a *differentiable function of* u, v and

$$\begin{aligned} dz &= (f_x g_u + f_y h_u)\, du + (f_x g_v + f_y h_v)\, dv \\ &= f_x(g_u\, du + g_v\, dv) + f_y(h_u\, du + h_v\, dv) \\ &= f_x\, dx + f_y\, dy \end{aligned}$$

$$\therefore \qquad dz = \frac{\partial z}{\partial x} dx + \frac{\partial z}{\partial y} dy.$$

Remark: The theorem establishes a fact of fundamental importance that *the first differential of a function is expressed always by the same formula, whether the variables concerned are independent or whether they are themselves functions of other independent variables.*

Note: The differential dz is sometimes referred to as the *total differential.*

7.1 Differentials of Higher Order of a Function of Functions

If $\frac{\partial z}{\partial x}, \frac{\partial z}{\partial y}$ are differentiable functions of x, y so that they are also differentiable functions of u, v, and dx, dy are differentiable functions of u, v, then from the preceding theorem we have

$$d^2 z = d(dz) = d\left(\frac{\partial z}{\partial x}\right)dx + \frac{\partial z}{\partial x} d^2 x + d\left(\frac{\partial z}{\partial y}\right)dy + \frac{\partial z}{\partial y} d^2 y$$

and on comparison with (2) and (3) of § 6, we see that

$$d^2 z = \frac{\partial^2 z}{\partial x^2} dx^2 + 2\frac{\partial^2 z}{\partial x \partial y} dx dy + \frac{\partial^2 z}{\partial y^2} dy^2 + \frac{\partial z}{\partial x} d^2 x + \frac{\partial z}{\partial y} d^2 y \qquad \text{...(1)}$$

$$= \left(\frac{\partial}{\partial x} dx + \frac{\partial}{\partial y} dy\right)^2 z + \frac{\partial z}{\partial x} d^2 x + \frac{\partial z}{\partial x} d^2 y \qquad \text{...(2)}$$

The differentials of higher orders can be written in the same manner, but their formation becomes more and more complicated and lengthy, and no simple general formula for $d^n z$ can be given.

The introduction of more than two *intermediary variables*[*] causes no fresh difficulty. Thus, when $z = f(x_1, x_2, x_3)$ and x_1, x_2, x_3 are not the independent variables,

$$d^2 z = \left(\frac{\partial}{\partial x_1} dx_1 + \frac{\partial}{\partial x_2} dx_2 + \frac{\partial}{\partial x_3} dx_3\right)^2 z + \frac{\partial z}{\partial x_1} d^2 x_1 + \frac{\partial z}{\partial x_2} d^2 x_2 + \frac{\partial z}{\partial x_3} d^2 x_3$$

Note: If x, y are linear functions of independent variables u and v, *i.e.*, x and y are of the form $x = a + bu + cv$, $y = a' + b'u + c'v$ then dx and dy are constants, and so d^2x, d^2y and all higher differentials of x and y are zero, and therefore

$$d^n z = \left(\frac{\partial}{\partial x} dx + \frac{\partial}{\partial y} dy\right)^n z,$$

the form being same as for independent x and y.

7.2 The Derivation of Composite Functions (*The chain rule*)

From the preceding theorem we deduce two important results:

I. *If*

(*i*) *x, y be differentiable functions of a single variable, and*

(*ii*) *z is differentiable function of x and y,*

then z possesses continuous derivative with respect to t, and

$$\frac{dz}{dt} = \frac{\partial z}{\partial x}\frac{dx}{dt} + \frac{\partial z}{\partial y}\frac{dy}{dt}$$

Because of (*i*),

$$dx = \frac{dx}{dt} \cdot dt, \text{ and } dy = \frac{dy}{dt} \cdot dt$$

Since z is a differentiable function of x and y, and x, y are differentiable functions of t, we deduce from § 7, that z is a differentiable function of t.

$$\therefore \qquad dz = \frac{dz}{dt} \cdot dt \qquad ...(1)$$

Also
$$dz = \frac{\partial z}{\partial x} dx + \frac{\partial z}{\partial y} dy = \frac{\partial z}{\partial x}\frac{dx}{dt} dt + \frac{\partial z}{\partial y}\frac{dy}{dt} dt \qquad ...(2)$$

From equations (1) and (2),

$$\frac{dz}{dt} = \frac{\partial z}{\partial x}\frac{dx}{dt} + \frac{\partial z}{\partial y}\frac{dy}{dt} \qquad ...(3)$$

[*] Variables like x, y which are functions of independent variables u, v are called *intermediary variables*.

Again because of conditions (*i*) and (*ii*), $\frac{dz}{dt}$ is a continuous function of t.

Corollary. *If $z = f(x, y)$ possesses nth order partial derivatives, and x, y are linear functions of a single variable t, i.e., $x = a + ht$, $y = b + kt$, where a, b, h, k are constants, then*

$$\frac{d^n z}{dt^n} = \left(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right)^n z$$

Now

$$\frac{dz}{dt} = \frac{\partial z}{\partial x}\frac{dx}{dt} + \frac{\partial z}{\partial y}\frac{dy}{dt} = h\frac{\partial z}{\partial x} + k\frac{\partial z}{\partial y} = \left(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right)z \qquad ...(1)$$

Replacing z by $\left(h\frac{\partial z}{\partial x} + k\frac{\partial z}{\partial y}\right)$ in (1), we get

$$\frac{d^2 z}{dt^2} = \frac{d}{dt}\left(\frac{dz}{dt}\right) = h\frac{\partial}{\partial x}\left(h\frac{\partial z}{\partial x} + k\frac{\partial z}{\partial y}\right) + k\frac{\partial}{\partial y}\left(h\frac{\partial z}{\partial x} + k\frac{\partial z}{\partial y}\right)$$

$$= h^2\frac{\partial^2 z}{\partial x^2} + 2hk\frac{\partial^2 z}{\partial x\partial y} + k^2\frac{\partial^2 z}{\partial y^2} = \left(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right)^2 z$$

By induction, we may obtain the required expression for $\frac{d^n z}{dt^n}$.

II. *If*

(*i*) *x, y are differentiable functions of two independent variables u and v, and*

(*ii*) *z is a differentiable function of x and y,*

then z possesses continuous partial derivatives with respect to u and v, and

$$\frac{\partial z}{\partial u} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial u} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial u}, \textit{ and } \frac{\partial z}{\partial v} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial v} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial v}$$

Because of (*i*)

$$\left.\begin{aligned} dx &= \frac{\partial x}{\partial u}du + \frac{\partial x}{\partial v}dv \\ dy &= \frac{\partial y}{\partial u}du + \frac{\partial y}{\partial v}dv \end{aligned}\right\} \qquad ...(1)$$

Since z is a differentiable function of x and y and x, y are differentiable functions of u and v, we deduce from § 7, that z is a differentiable function of u, and v, and

$$dz = \frac{\partial z}{\partial u}du + \frac{\partial z}{\partial v}dv \qquad ...(2)$$

Also $$dz = \frac{\partial z}{\partial x}dx + \frac{\partial z}{\partial y}dy = \frac{\partial z}{\partial x}\left(\frac{\partial z}{\partial u}du + \frac{\partial z}{\partial v}dv\right) + \frac{\partial z}{\partial y}\left(\frac{\partial y}{\partial u}du + \frac{\partial y}{\partial v}dv\right)$$

$$= \left(\frac{\partial z}{\partial x}\frac{\partial x}{\partial u} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial u}\right)du + \left(\frac{\partial z}{\partial x}\frac{\partial x}{\partial v} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial v}\right)dv \qquad ...(3)$$

Hence, from equations (2) and (3), we get

$$\frac{\partial z}{\partial u} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial u} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial u}, \text{ and } \frac{\partial z}{\partial v} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial v} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial v}$$

Again, because of conditions (*i*) and (*ii*) we see that $\frac{\partial z}{\partial u}$ and $\frac{\partial z}{\partial v}$ are continuous functions of u, v.

Note: In (1) when x is a function of a single variable t, we have $dx = \frac{dx}{dt}dt$, so that the derivative $\frac{dx}{dt}$ appears as the coefficient of a differential and that is precisely the reason why the derivative is also called the *differential coefficient*.

Example 19. If $z = e^{xy^2}$, $x = t\cos t$, $y = t\sin t$, compute $\frac{dz}{dt}$ at $t = \frac{\pi}{2}$.

■ $$\frac{dz}{dt} = \frac{\partial z}{\partial x}\frac{dx}{dt} + \frac{\partial z}{\partial y}\frac{dy}{dt} = (y^2 e^{xy^2})(\cos t - t\sin t) + (2xye^{xy^2})(\sin t + t\cos t)$$

At $t = \frac{\pi}{2} \Rightarrow x = 0,\ y = \frac{\pi}{2}$.

$$\therefore \qquad \left[\frac{dz}{dt}\right]_{t=\pi/2} = \frac{\pi^2}{4}\left(-\frac{\pi}{2}\right) = -\frac{\pi^3}{8}.$$

Example 20. If $z = x^3 - xy + y^3$, $x = r\cos\theta$, $y = r\sin\theta$, find $\frac{\partial z}{\partial r}, \frac{\partial z}{\partial \theta}$.

■ $$\frac{\partial z}{\partial r} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial r} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial r} = (3x^2 - y)\cos\theta + (3y^2 - x)\sin\theta$$

$$\frac{\partial z}{\partial \theta} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial \theta} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial \theta} = (3x^2 - y)(-r\sin\theta) + (3y^2 - x)\, r\cos\theta.$$

Example 21. Show that $z = f(x^2 y)$, where f is differentiable, satisfies

$$x\left(\frac{\partial z}{\partial x}\right) = 2y\left(\frac{\partial z}{\partial y}\right).$$

■ Let $x^2 y = u$, so that $z = f(u)$. Thus

$$\frac{\partial z}{\partial x} = \frac{\partial z}{\partial u}\frac{\partial u}{\partial x} = f'(u) \cdot 2xy$$

$$\frac{\partial z}{\partial y} = \frac{\partial z}{\partial u}\frac{\partial u}{\partial y} = f'(u) \cdot x^2$$

$\therefore$
$$x\frac{\partial z}{\partial x} = f'(u)\, 2x^2 y = 2y\frac{\partial z}{\partial y}$$

Aliter. $dz = f'(u)\,du = f'(x^2 y)\,(2xy\,dx + x^2\,dy)$

Also $dz = \dfrac{\partial z}{\partial x}dx + \dfrac{\partial z}{\partial x}dy$

Then $\dfrac{\partial z}{\partial x} = 2xyf'(x^2 y),\ \dfrac{\partial z}{\partial y} = x^2 f'(x^2 y)$

The result now follows as above.

Example 22. If for all values of the parameter λ, and for some constant n, $F(\lambda x, \lambda y) = \lambda^n F(x, y)$ (F is then called a *homogeneous function* of degree n), identically where F is assumed differentiable, prove that $x\dfrac{\partial F}{\partial x} + y\dfrac{\partial F}{\partial y} = nF$. Hence show that, for $F(x, y) = x^4 y^2 \sin^{-1}\dfrac{y}{x}$,

$$x\frac{\partial F}{\partial x} + y\frac{\partial F}{\partial y} = 6F.$$

■ Let $\lambda x = u$, $\lambda y = v$. Then

$$F(u, v) = \lambda^n F(x, y) \qquad \text{...(1)}$$

The derivative w.r.t. λ of the left side of (1) is

$$\frac{\partial F}{\partial \lambda} = \frac{\partial F}{\partial u}\frac{\partial u}{\partial \lambda} + \frac{\partial F}{\partial v}\frac{\partial v}{\partial \lambda} = x\frac{\partial F}{\partial u} + y\frac{\partial F}{\partial v}$$

The derivative w.r.t. λ of the right side of (1) is $n\lambda^{n-1} F(x, y)$. Then

$$x\frac{\partial F}{\partial u} + y\frac{\partial F}{\partial v} = n\lambda^{n-1} F$$

The result follows for $\lambda = 1$, then $u = x$, $v = y$.

Again, since $F(\lambda x, \lambda y) = (\lambda x)^4 (\lambda y)^2 \sin^{-1} y/x = \lambda^6 F(x, y)$, the result follows for $n = 6$.

That it is so, can also be shown by direct differentiation.

Example 23. If z is given as a function of two independent variables x and y, change the variables so that x becomes the function, and z and y the independent variables, and express the first and second order partial derivatives of x with respect to z and y in terms of the derivatives of z with respect to x and y.

- When x and y are independent variables and z the dependent, a usual notation (which will be often employed) is

$$\frac{\partial z}{\partial x} = p, \frac{\partial z}{\partial y} = q, \frac{\partial^2 z}{\partial x^2} = r, \frac{\partial^2 z}{\partial x \partial y} = s, \frac{\partial^2 z}{\partial y^2} = t$$

We know

$$dz = \frac{\partial z}{\partial x} dx + \frac{\partial z}{\partial y} dy = p\,dx + q\,dy \qquad ...(1)$$

Again, when z and y are independent and x the function,

$$dx = \frac{\partial x}{\partial z} dz + \frac{\partial x}{\partial y} dy \qquad ...(2)$$

From equation (1),

$$dx = \frac{1}{p} dz - \frac{q}{p} dy \qquad ...(3)$$

Comparing the coefficients in equation (2) and equation (3), we get

$$\frac{\partial x}{\partial z} = \frac{1}{p}, \quad \frac{\partial x}{\partial y} = -\frac{q}{p} \qquad ...(4)$$

Taking the differential of the first, we have

$$d\left(\frac{\partial x}{\partial z}\right) = d\left(\frac{1}{p}\right)$$

or

$$\frac{\partial^2 x}{\partial z^2} dz + \frac{\partial^2 x}{\partial y \partial z} dy = -\frac{1}{p^2} dp \qquad ...(5)$$

But p is a function of x and y.

$\therefore$
$$dp = \frac{\partial p}{\partial x} dx + \frac{\partial p}{\partial y} dy = r\,dx + s\,dy$$

Hence from equation (5),

$$\frac{\partial^2 x}{\partial z^2} dz + \frac{\partial^2 x}{\partial y \partial z} dy = -\frac{1}{p^2}(r\ dx + s\ dy) = -\frac{r}{p^3} dz + \frac{rq - sp}{p^3} dy \qquad \text{[using (3)]}$$

In this equation, we have only the differentials of independent variables and can therefore equate the coefficients of dz and dy, hence

$$\frac{\partial^2 x}{\partial z^2} = -\frac{r}{p^3}, \frac{\partial^2 x}{\partial y \partial z} = \frac{rq - sp}{p^3} \qquad ...(6)$$

In the same way, the second equation of (4) gives

$$\frac{\partial^2 x}{\partial z \partial y} dz + \frac{\partial^2 x}{\partial y^2} dy = -\frac{p\,dq - q\,dp}{p^2} \qquad ...(7)$$

But $dq = \frac{\partial q}{\partial x} dx + \frac{\partial q}{\partial y} dy = s\,dx + t\,dy$, substituting, as before for dp, dq and using (3), we have from (7)

$$\frac{\partial^2 x}{\partial z \partial y} dz + \frac{\partial^2 x}{\partial y^2} dy = \frac{rq - sp}{p^3} dz + \frac{2pqs - tp^2 - rq^2}{p^3} dy$$

and, therefore

$$\frac{\partial^2 x}{\partial y^2} = \frac{2pqs - tp^2 - rq^2}{p^3}$$

The value of $\partial^2 x/\partial z \partial y$ being the same as $\partial^2 x/\partial y \partial z$.

Note: $$\frac{\partial^2 x}{\partial z^2} \cdot \frac{\partial^2 x}{\partial y^2} - \left(\frac{\partial^2 x}{\partial z \partial y}\right)^2 = \frac{rt - s^2}{p^4}$$

8. CHANGE OF VARIABLES

In problems involving change of variables it is frequently required to transform a particular expression involving a combination of derivatives with respect to a set of variables, in terms of derivatives with respect to another set of variables. A general method, illustrating the principles involved is given below, but it can often be modified so as to reduce the algebraic work.

We shall consider derivatives up to second order only. The higher derivatives may be obtained by exactly the same method; fortunately they are not often required. The algebra of the transformation is tedious but the method seems simple.

Problem. *If z is a function $f(x, y)$ of the independent variables x, y, and if x, y are changed to new independent variables u, v by the substitutions $x = \phi(u, v)$, $y = \psi(u, v)$, it is required to express the derivatives of z with respect to x, y in terms of u, v and the derivatives of z with respect to u, v.*

It is understood that f, ϕ, ψ are differentiable (or possesses continuous partial derivatives) with respect to the corresponding variables.

But rule II of § 7.2, we have

$$\left.\begin{aligned} \frac{\partial z}{\partial u} &= \frac{\partial z}{\partial x}\frac{\partial x}{\partial u} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial u} \\ \frac{\partial z}{\partial v} &= \frac{\partial z}{\partial x}\frac{\partial x}{\partial v} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial v} \end{aligned}\right\} \qquad ...(1)$$

Solving these for $\dfrac{\partial z}{\partial x}$ and $\dfrac{\partial z}{\partial y}$, we get

$$\left.\begin{aligned}\frac{\partial z}{\partial x} &= A\frac{\partial z}{\partial u} + B\frac{\partial z}{\partial v}\\ \frac{\partial z}{\partial y} &= C\frac{\partial z}{\partial u} + D\frac{\partial z}{\partial v}\end{aligned}\right\} \qquad ...(2)$$

where $$A = \frac{\partial y}{\partial v}\Big/J,\ B = -\frac{\partial y}{\partial u}\Big/J,\ C = -\frac{\partial x}{\partial v}\Big/J,\ D = \frac{\partial x}{\partial u}\Big/J \text{ and}$$

$$J = \frac{\partial(x, y)}{\partial(u, v)} = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix}, \text{ called the } \textit{Jacobian},$$

are functions of u and v. (Refer to § 2 of Chapter 16, for properties of Jacobians.)

Thus,

$$\frac{\partial z}{\partial x} = -\frac{\partial(y, z)}{\partial(u, v)}\Big/\frac{\partial(x, y)}{\partial(u, v)} \text{ and } \frac{\partial z}{\partial y} = -\frac{\partial(z, x)}{\partial(u, v)}\Big/\frac{\partial(x, y)}{\partial(u, v)}$$

Equation (2) expresses $\dfrac{\partial z}{\partial x}$ and $\dfrac{\partial z}{\partial y}$ in terms of $A, B, C, D, \dfrac{\partial z}{\partial u}, \dfrac{\partial z}{\partial v}$ which are all functions of u, v, and not contain x, y explicitly.

From (2),

$$\frac{\partial}{\partial x} = A\frac{\partial}{\partial u} + B\frac{\partial}{\partial v},\quad \frac{\partial}{\partial y} = C\frac{\partial}{\partial u} + D\frac{\partial}{\partial v}$$

Replacing z by $\dfrac{\partial z}{\partial x}$ in equation (2), we get

$$\begin{aligned}\frac{\partial^2 z}{\partial x^2} &= \frac{\partial}{\partial x}\left(\frac{\partial z}{\partial x}\right) = \frac{\partial}{\partial x}\left(A\frac{\partial z}{\partial u} + B\frac{\partial z}{\partial v}\right) = \left(A\frac{\partial}{\partial u} + B\frac{\partial}{\partial v}\right)\left(A\frac{\partial z}{\partial u} + B\frac{\partial z}{\partial v}\right)\\ &= A\frac{\partial}{\partial u}\left(A\frac{\partial z}{\partial u} + B\frac{\partial z}{\partial v}\right) + B\frac{\partial}{\partial v}\left(A\frac{\partial z}{\partial u} + B\frac{\partial z}{\partial v}\right)\\ &= A^2\frac{\partial^2 z}{\partial u^2} + 2AB\frac{\partial^2 z}{\partial u\partial v} + B^2\frac{\partial^2 z}{\partial v^2} + \left(A\frac{\partial A}{\partial u} + B\frac{\partial A}{\partial v}\right)\frac{\partial z}{\partial u} + \left(A\frac{\partial B}{\partial u} + B\frac{\partial B}{\partial v}\right)\frac{\partial z}{\partial v}\end{aligned}$$

The values of $\dfrac{\partial^2 z}{\partial y^2}$ and $\dfrac{\partial^2 z}{\partial x\partial y}$ may be found in the same way.

Remark: The expression for $\frac{\partial z}{\partial x}, \frac{\partial z}{\partial y}$ can also be found as follows:

$$\frac{\partial z}{\partial x}=\frac{\partial z}{\partial u}\frac{\partial u}{\partial x}+\frac{\partial z}{\partial v}\frac{\partial v}{\partial x}, \frac{\partial z}{\partial y}=\frac{\partial z}{\partial u}\frac{\partial u}{\partial y}+\frac{\partial z}{\partial v}\frac{\partial v}{\partial y} \quad ...(3)$$

To obtain $\frac{\partial u}{\partial x}$, differentiate $x=\phi(u, v)$ and $y=\psi(u, v)$ with respect to x,

$$\therefore \quad 1=\frac{\partial x}{\partial u}\frac{\partial u}{\partial x}+\frac{\partial x}{\partial v}\frac{\partial v}{\partial x}, \text{ and } 0=\frac{\partial y}{\partial u}\frac{\partial u}{\partial x}+\frac{\partial y}{\partial v}\frac{\partial v}{\partial x}$$

and these give

$$\frac{\partial u}{\partial x}=A, \frac{\partial v}{\partial x}=B \quad ...(4)$$

Differentiating $x=\phi(u, v)$, $y=\psi(u, v)$ with respect to y, we get

$$\frac{\partial u}{\partial y}=C, \frac{\partial v}{\partial y}=D \quad ...(5)$$

Equation (3) now gives the required result.

Note: In the 'change of variables' the variables x, y which are functions of u, v are called the *Intermediate variables*, while u, v are independent variables.

Example 24. If $u = F(x, y, z)$, and $z = f(x, y)$, find a formula for $\partial^2 u/\partial x^2$ in terms of the derivatives of F and the derivatives of z.

■ In the expression for F we consider x, y, z as intermediate variables, while in the expression for f we consider x and y as independent variables.

Now
$$\frac{\partial u}{\partial x}=F_x\frac{\partial x}{\partial x}+F_y\frac{\partial y}{\partial x}+F_z\frac{\partial z}{\partial x},$$

but since x and y are independent, $\frac{\partial y}{\partial x}=0$.

Also,
$$\frac{\partial x}{\partial x}=1$$

$$\therefore \quad \frac{\partial u}{\partial x}=F_x+F_z\frac{\partial z}{\partial x}$$

Differentiating a second time, we get

$$\frac{\partial^2 u}{\partial x^2}=F_{xx}\frac{\partial x}{\partial x}+F_{yx}\frac{\partial y}{\partial x}+F_{zx}\frac{\partial z}{\partial x}+\frac{\partial z}{\partial x}\left(F_{xz}\frac{\partial x}{\partial x}+F_{yz}\frac{\partial y}{\partial x}+F_{zz}\frac{\partial z}{\partial x}\right)+F_z\frac{\partial^2 z}{\partial x^2}$$

$$=F_{x^2}+2F_{zx}\frac{\partial z}{\partial x}+F_{z^2}\left(\frac{\partial z}{\partial x}\right)^2+F_z\frac{\partial^2 z}{\partial x^2}$$

of course, F and f are supposed to be differentiable.

Example 25. Show that $f(xy, z-2x) = 0$ satisfies, under suitable conditions, the equation $x\frac{\partial z}{\partial x} - y\frac{\partial z}{\partial y} = 2x$. What are these conditions?

■ Let $u = xy$, $v = z - 2x$; then $f(u, v) = 0$, and

$$df = f_u\, du + f_v\, dv = f_u(x\, dy + y\, dx) + f_v(dz - 2\, dx) = 0$$

Taking z as dependent variable and x and y as independent variables, we have

$$dz = \frac{\partial z}{\partial x}dx + \frac{\partial z}{\partial y}dy$$

$\therefore$
$$df = f_u(x\,dy + y\,dx) + f_v\left(\frac{\partial z}{\partial x}dx + \frac{\partial z}{\partial y}dy - 2\,dx\right) = 0$$

or

$$\left\{yf_u + f_v\left(\frac{\partial z}{\partial x} - 2\right)\right\}dx + \left\{xf_u + f_v\frac{\partial z}{\partial y}\right\}dy = 0$$

But, since x and y are independent, we have

$$yf_u + f_v\left(\frac{\partial z}{\partial x} - 2\right) = 0, \text{ and } xf_u + f_v\frac{\partial z}{\partial y} = 0.$$

Finding f_u from one equation and putting in the other, we get

$$x\frac{\partial z}{\partial x} - y\frac{\partial z}{\partial y} = 2x, \text{ provided } f_v \neq 0$$

Thus, the result holds when f is differentiable and $f_v \neq 0$ (and then $f_u \neq 0$).

Example 26. Prove that, by the transformations $u = x - ct$, $v = x + ct$, the partial differential equation

$$\frac{\partial^2 z}{\partial t^2} = c^2\frac{\partial^2 z}{\partial x^2} \text{ reduces to } \frac{\partial^2 z}{\partial u\partial v} = 0$$

■ In this problem, we consider z as a function of u and v which are linear functions of the independent variables x and t.

Now by § 7.1,

$$d^2z = \frac{\partial^2 z}{\partial u^2}du^2 + 2\frac{\partial^2 z}{\partial u\partial v}du\,dv + \frac{\partial^2 z}{\partial v^2}dv^2 + \frac{\partial z}{\partial u}d^2u + \frac{\partial z}{\partial v}d^2v$$

But $u = x - ct$, and $v = x + ct$

$\therefore$ $du = dx - cdt,\ dv = dx + cdt$

$d^2u = 0,\ d^2v = 0$

$\therefore$
$$d^2z = \frac{\partial^2 z}{\partial u^2}(dx - cdt)^2 + 2\frac{\partial^2 z}{\partial u\partial v}(dx - cdt)\,(dx + cdt) + \frac{\partial^2 z}{\partial v^2}(dx + cdt)^2$$

$$= \left(\frac{\partial^2 z}{\partial u^2} + 2\frac{\partial^2 z}{\partial u \partial v} + \frac{\partial^2 z}{\partial v^2} \right) dx^2 + B\ dx\ dt + c^2 \left\{ \frac{\partial^2 z}{\partial u^2} - 2\frac{\partial^2 z}{\partial u \partial v} + \frac{\partial^2 z}{\partial v^2} \right\} dt^2 \qquad \text{...(1)}$$

the coefficient of $dx\ dt$ is written as B since its actual value is not required for this problem.

Again regarding z as a function of independent variables x and t, we have (by § 6)

$$d^2 z = \frac{\partial^2 z}{\partial x^2} dx^2 + 2\frac{\partial^2 z}{\partial x \partial t} dx\, dt + \frac{\partial^2 z}{\partial t^2} dt^2 \qquad \text{...(2)}$$

From equations (1) and (2) by comparing the coefficients

$$\frac{\partial^2 z}{\partial x^2} = \frac{\partial^2 z}{\partial u^2} + 2\frac{\partial^2 z}{\partial u \partial v} + \frac{\partial^2 z}{\partial v^2} \qquad \text{...(3)}$$

$$\frac{\partial^2 z}{\partial t^2} = c^2 \left\{ \frac{\partial^2 z}{\partial u^2} - 2\frac{\partial^2 z}{\partial u \partial v} + \frac{\partial^2 z}{\partial v^2} \right\} \qquad \text{...(4)}$$

$$\therefore \qquad \frac{\partial^2 z}{\partial t^2} = c^2 \frac{\partial^2 z}{\partial x^2} \Rightarrow \frac{\partial^2 z}{\partial u \partial v} = 0$$

Note: $\dfrac{\partial^2 z}{\partial t^2} + c^2 \dfrac{\partial^2 z}{\partial x^2}$ reduces to $2c^2 \left(\dfrac{\partial^2 z}{\partial u^2} + \dfrac{\partial^2 z}{\partial v^2} \right)$.

Example 27. Prove that $\dfrac{\partial^2 V}{\partial x^2} + \dfrac{\partial^2 V}{\partial y^2}$ is invariant for change of rectangular axes.

■ The change of origin does not affect the expression, for then $x = a + x'$, $y = b + y'$, and $dx = dx'$, $dy = dy'$, and all the partial derivatives V remain of the same form.

Let the axes turn through an angle α, so that

$$x = x' \cos\alpha - y' \sin\alpha, \text{ and } y = x' \sin\alpha + y' \cos\alpha,$$

where α is a constant.

We have

$$\begin{aligned} dV &= \frac{\partial V}{\partial x} dx + \frac{\partial V}{\partial y} dy \\ &= \frac{\partial V}{\partial x}(dx' \cos\alpha - dy' \sin\alpha) + \frac{\partial V}{\partial y}(dx' \sin\alpha + dy' \cos\alpha) \\ &= \left(\frac{\partial V}{\partial x} \cos\alpha + \frac{\partial V}{\partial y} \sin\alpha \right) dx' + \left(-\frac{\partial V}{\partial x} \sin\alpha + \frac{\partial V}{\partial y} \sin\alpha \right) dy' \end{aligned}$$

Also

$$dV = \frac{\partial V}{\partial x'} dx' + \frac{\partial V}{\partial y'} dy'$$

$$\therefore \qquad \frac{\partial V}{\partial x'} = \frac{\partial V}{\partial x}\cos\alpha + \frac{\partial V}{\partial y}\sin\alpha \text{ and } \frac{\partial V}{\partial y'} = -\frac{\partial V}{\partial x}\sin\alpha + \frac{\partial V}{\partial y}\cos\alpha \qquad ...(1)$$

These give ...(2)

$$\left.\begin{aligned} \frac{\partial V}{\partial x} &= \frac{\partial V}{\partial x'}\cos\alpha - \frac{\partial V}{\partial y'}\sin\alpha \\ \frac{\partial V}{\partial y} &= \frac{\partial V}{\partial x'}\sin\alpha + \frac{\partial V}{\partial y'}\cos\alpha \end{aligned}\right\}$$

$$\therefore \qquad \frac{\partial^2 V}{\partial x^2} = \frac{\partial}{\partial x}\left(\frac{\partial V}{\partial x}\right) = \left(\frac{\partial}{\partial x'}\cos\alpha - \frac{\partial}{\partial y'}\sin\alpha\right)\left(\frac{\partial V}{\partial x'}\cos\alpha - \frac{\partial V}{\partial y'}\sin\alpha\right)$$

$$= \frac{\partial^2 V}{\partial x'^2}\cos^2\alpha - 2\frac{\partial^2 V}{\partial x'\partial y'}\sin\alpha\cos\alpha + \frac{\partial^2 V}{\partial y'^2}\sin^2\alpha$$

$$\frac{\partial^2 V}{\partial y^2} = \frac{\partial}{\partial y}\left(\frac{\partial V}{\partial y}\right) = \left(\frac{\partial}{\partial x'}\sin\alpha + \frac{\partial}{\partial y'}\cos\alpha\right)\left(\frac{\partial V}{\partial x'}\sin\alpha + \frac{\partial V}{\partial y'}\cos\alpha\right)$$

$$= \frac{\partial^2 V}{\partial x'^2}\sin^2\alpha + 2\frac{\partial^2 V}{\partial x'\partial y'}\sin\alpha\cos\alpha + \frac{\partial^2 V}{\partial y'^2}\cos^2\alpha$$

$$\therefore \qquad \frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} = \frac{\partial^2 V}{\partial x'^2} + \frac{\partial^2 V}{\partial y'^2}$$

Thus, the expression remains invariant.

Note: If we write $x' = x\cos\alpha + y\sin\alpha$, $y' = -x\sin\alpha + y\cos\alpha$ the procedure is slightly simplified.

Example 28. If V is a function of two variables x and y and $x = r\cos\theta$, $y = r\sin\theta$, prove that

$$\frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} = \frac{\partial^2 V}{\partial r^2} + \frac{1}{r^2}\frac{\partial^2 V}{\partial \theta^2} + \frac{1}{r}\frac{\partial V}{\partial r}$$

■ We have

$$\frac{\partial V}{\partial r} = \frac{\partial V}{\partial x}\frac{\partial x}{\partial r} + \frac{\partial V}{\partial y}\frac{\partial y}{\partial r} = \cos\theta\frac{\partial V}{\partial x} + \sin\theta\frac{\partial V}{\partial y}$$

$$\frac{\partial V}{\partial \theta} = \frac{\partial V}{\partial x}\frac{\partial x}{\partial \theta} + \frac{\partial V}{\partial y}\frac{\partial y}{\partial \theta} = -r\sin\theta\frac{\partial V}{\partial x} + r\cos\theta\frac{\partial V}{\partial y}$$

Solving these, we get

$$\frac{\partial V}{\partial x} = \cos\theta\frac{\partial V}{\partial r} - \frac{\sin\theta}{r}\frac{\partial V}{\partial \theta} \Rightarrow \frac{\partial}{\partial x} = \cos\theta\frac{\partial}{\partial r} - \frac{\sin\theta}{r}\frac{\partial}{\partial \theta}$$

CHAPTER 15

$$\frac{\partial V}{\partial y} = \sin\theta \frac{\partial V}{\partial r} + \frac{\cos\theta}{r}\frac{\partial V}{\partial y} \Rightarrow \frac{\partial}{\partial y} = \sin\theta \frac{\partial}{\partial r} + \frac{\cos\theta}{r}\frac{\partial}{\partial \theta}$$

Hence,

$$\frac{\partial^2 V}{\partial x^2} = \frac{\partial}{\partial x}\left(\frac{\partial V}{\partial x}\right)$$

$$= \left(\cos\theta \frac{\partial}{\partial r} - \frac{\sin\theta}{r}\frac{\partial}{\partial\theta}\right)\left(\cos\theta\frac{\partial V}{\partial r} - \frac{\sin\theta}{r}\frac{\partial V}{\partial\theta}\right)$$

$$= \cos^2\theta \frac{\partial^2 V}{\partial r^2} - \frac{2\sin\theta\cos\theta}{r}\frac{\partial^2 V}{\partial r\partial\theta} + \frac{\sin^2\theta}{r^2}\frac{\partial^2 V}{\partial\theta^2} + \frac{2\cos\theta\sin\theta}{r^2}\frac{\partial V}{\partial\theta} + \frac{\sin^2\theta}{r}\frac{\partial V}{\partial r}$$

Similarly,

$$\frac{\partial^2 V}{\partial y^2} = \sin^2\theta \frac{\partial^2 V}{\partial r^2} + \frac{2\sin\theta\cos\theta}{r}\cdot\frac{\partial^2 V}{\partial r\partial\theta} + \frac{\cos^2\theta}{r^2}\frac{\partial^2 V}{\partial\theta^2} - \frac{2\cos\theta\sin\theta}{r^2}\frac{\partial V}{\partial\theta} + \frac{\cos^2\theta}{r}\frac{\partial V}{\partial r}$$

$$\therefore \quad \frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} = \frac{\partial^2 V}{\partial r^2} + \frac{1}{r^2}\frac{\partial^2 V}{\partial\theta^2} + \frac{1}{r}\frac{\partial V}{\partial r}$$

Deduction 1.

$$\frac{\partial^2 V}{\partial x\partial y} = \left(\cos\theta\frac{\partial}{\partial r} - \frac{\sin\theta}{r}\frac{\partial}{\partial\theta}\right)\left(\sin\theta\frac{\partial V}{\partial r} + \frac{\cos\theta}{r}\frac{\partial V}{\partial\theta}\right)$$

$$= \cos\theta\sin\theta\frac{\partial^2 V}{\partial r^2} + \frac{\cos 2\theta}{r}\frac{\partial^2 V}{\partial r\partial\theta} - \frac{\sin\theta\cos\theta}{r^2}\frac{\partial^2 V}{\partial\theta^2} - \frac{\cos 2\theta}{r^2}\frac{\partial V}{\partial\theta} - \frac{\sin\theta\cos\theta}{r}\frac{\partial V}{\partial r}$$

Also

$$x^2 - y^2 = r^2\cos 2\theta, \text{ and } 4xy = 2r^2\sin 2\theta$$

$$\therefore \quad (x^2 - y^2)\left(\frac{\partial^2 V}{\partial x^2} - \frac{\partial^2 V}{\partial y^2}\right) + 4\,xy\frac{\partial^2 V}{\partial x\partial y} = r^2\frac{\partial^2 V}{\partial r^2} - r\frac{\partial V}{\partial r} - \frac{\partial^2 V}{\partial\theta^2}$$

Deduction 2. To show that $\dfrac{\partial^2\theta}{\partial x\partial y} = -\dfrac{\cos 2\theta}{r^2}$.

Here $x = r\cos\theta$, $y = r\sin\theta$

Differentiating w.r.t. r, θ are (functions of x and y)

$$\left.\begin{aligned} 1 &= \frac{\partial r}{\partial x}\cos\theta - r\sin\theta\frac{\partial\theta}{\partial x} \\ 0 &= \frac{\partial r}{\partial x}\sin\theta - r\cos\theta\frac{\partial\theta}{\partial x}\end{aligned}\right\} \quad \begin{aligned} \therefore\ \frac{\partial r}{\partial x} &= \cos\theta \\ \frac{\partial\theta}{\partial x} &= -\frac{\sin\theta}{r}\end{aligned}$$

Differentiating w.r.t. y,

$$\left.\begin{aligned} 0 &= \frac{\partial r}{\partial y}\cos\theta - r\sin\theta\frac{\partial \theta}{\partial y} \\ 1 &= \frac{\partial r}{\partial y}\sin\theta - r\cos\theta\frac{\partial \theta}{\partial y}\end{aligned}\right\} \quad \therefore \quad \begin{aligned}\frac{\partial r}{\partial y} &= \sin\theta \\ \frac{\partial \theta}{\partial y} &= \frac{\cos\theta}{r}\end{aligned}$$

Now $\dfrac{\partial^2\theta}{\partial x\partial y} = \dfrac{\partial}{\partial x}\left(\dfrac{\partial \theta}{\partial y}\right) = \dfrac{\partial}{\partial x}\left(\dfrac{\cos\theta}{r}\right) = -\dfrac{\cos\theta}{r^2}\dfrac{\partial r}{\partial x} - \dfrac{\sin\theta}{r}\dfrac{\partial \theta}{\partial x} = -\dfrac{\cos 2\theta}{r^2}$

Similarly, it may be shown that

$$\frac{\partial^2\theta}{\partial y\partial x} = -\frac{\cos 2\theta}{r^2}$$

$$\frac{\partial^2 r}{\partial x\partial y} = -\frac{\sin\theta\cos\theta}{r} = \frac{\partial^2 r}{\partial y\partial x}$$

Note: Using the method of Ded. 2, we get

$$\frac{\partial V}{\partial x} = \frac{\partial V}{\partial r}\frac{\partial r}{\partial x} + \frac{\partial V}{\partial \theta}\frac{\partial \theta}{\partial x} = \cos\theta\frac{\partial V}{\partial r} - \frac{\sin\theta}{r}\frac{\partial V}{\partial \theta}$$

and

$$\frac{\partial V}{\partial y} = \sin\theta\frac{\partial V}{\partial r} + \frac{\cos\theta}{r}\frac{\partial V}{\partial \theta}$$

which are same as in the above example.

Example 29. Given that F is a function of x and y and that $x = e^u + e^{-v}$, $y = e^v + e^{-u}$, prove that

$$\frac{\partial^2 F}{\partial u^2} - 2\frac{\partial^2 F}{\partial u\partial v} + \frac{\partial^2 F}{\partial v^2} = x^2\frac{\partial^2 F}{\partial x^2} - 2xy\frac{\partial^2 F}{\partial x\partial y} + y^2\frac{\partial^2 F}{\partial y^2} + x\frac{\partial F}{\partial x} + y\frac{\partial F}{\partial y}$$

■ Here F is a function of the intermediary variables x and y; and u, v are the independent variables.
Now $x = e^u + e^{-v}$, and $y = e^v + e^{-u}$

$$\therefore \quad dx = e^u\,du - e^{-v}\,dv, \qquad dy = e^v\,dv - e^{-u}\,du$$

$$d^2x = e^u\,du^2 + e^{-v}\,dv^2, \qquad d^2y = e^v\,dv^2 + e^{-u}\,du^2$$

$$d^2F = \frac{\partial^2 F}{\partial x^2}dx^2 + 2\frac{\partial^2 F}{\partial x\partial y}dx\,dy + \frac{\partial^2 F}{\partial y^2}dy^2 + \frac{\partial F}{\partial x}d^2x + \frac{\partial F}{\partial y}d^2y$$

$$= \frac{\partial^2 F}{\partial x^2}(e^u\,du - e^{-v}\,dv)^2 + 2\frac{\partial^2 F}{\partial x\partial y}(e^u\,du - e^{-v}\,dv)\,(e^v\,dv - e^{-u}\,du)$$

$$+ (e^v\,dv - e^{-u}\,du)^2\frac{\partial^2 F}{\partial y^2} + \frac{\partial F}{\partial x}(e^u\,du^2 + e^{-v}\,dv^2) + \frac{\partial F}{\partial y}(e^v\,dv^2 + e^{-u}\,du^2)$$

$$= \left(e^{2u} \frac{\partial^2 F}{\partial x^2} - 2 \frac{\partial^2 F}{\partial x \partial y} + e^{-2u} \frac{\partial^2 F}{\partial y^2} + e^{u} \frac{\partial F}{\partial x} + e^{-u} \frac{\partial F}{\partial y} \right) du^2$$

$$+ 2 \left[-e^{u} e^{-v} \frac{\partial^2 F}{\partial x^2} + (e^{u} e^{v} + e^{-u} e^{-v}) \frac{\partial^2 F}{\partial x \partial y} - e^{v} e^{-u} \frac{\partial F}{\partial y^2} \right] du\, dv$$

$$+ \left(e^{-2v} \frac{\partial^2 F}{\partial x^2} - 2 \frac{\partial^2 F}{\partial x \partial y} + e^{2v} \frac{\partial^2 F}{\partial y^2} + e^{-v} \frac{\partial F}{\partial x} + e^{v} \frac{\partial F}{\partial y} \right) dv^2$$

Also $d^2 F = \dfrac{\partial^2 F}{\partial u^2} du^2 + 2 \dfrac{\partial^2 F}{\partial u \partial v} du\, dv + \dfrac{\partial^2 F}{\partial v^2} dv^2$

Comparing the coefficients, we get

$$\frac{\partial^2 F}{\partial u^2} = e^{2u} \frac{\partial^2 F}{\partial x^2} - 2 \frac{\partial^2 F}{\partial x \partial y} + e^{-2u} \frac{\partial^2 F}{\partial y^2} + e^{u} \frac{\partial F}{\partial x} + e^{-u} \frac{\partial F}{\partial y}$$

$$\frac{\partial^2 F}{\partial u \partial v} = -e^{-u} e^{-v} \frac{\partial^2 F}{\partial x^2} + (e^{u} e^{v} + e^{-u} e^{-v}) \frac{\partial^2 F}{\partial x \partial y} - e^{v} e^{-u} \frac{\partial^2 F}{\partial y^2}$$

$$\frac{\partial^2 F}{\partial v^2} = e^{-2v} \frac{\partial^2 F}{\partial x^2} - 2 \frac{\partial^2 F}{\partial x \partial y} + e^{2v} \frac{\partial^2 F}{\partial y^2} + e^{-v} \frac{\partial F}{\partial x} + e^{v} \frac{\partial F}{\partial y}$$

$$\therefore \quad \frac{\partial^2 F}{\partial u^2} - 2 \frac{\partial^2 F}{\partial u \partial v} + \frac{\partial^2 F}{\partial x^2} = x^2 \frac{\partial^2 F}{\partial x^2} - 2xy \frac{\partial^2 F}{\partial x \partial y} + y^2 \frac{\partial^2 F}{\partial y^2} + x \frac{\partial F}{\partial x} + y \frac{\partial F}{\partial y}$$

Remark: The four examples 26 to 29 give different methods for solving problems of this type. However any one method could be used to solve all such problems. The reader is advised to try.

EXERCISE

1. If $x = u \cos\alpha - v \sin\alpha$ and $y = u \sin\alpha + v \cos\alpha$, where α is a constant, show that

$$\left(\frac{\partial V}{\partial x} \right)^2 + \left(\frac{\partial V}{\partial y} \right)^2 = \left(\frac{\partial V}{\partial u} \right)^2 + \left(\frac{\partial V}{\partial v} \right)^2$$

2. If $2axz + 2byz + cz^2 = k$, $ax + by + cz = R$, prove that

$$R^3 \frac{\partial^2 z}{\partial x^2} = a^2 k,\ R^3 \frac{\partial^2 z}{\partial x \partial y} = abk,\ R^3 \frac{\partial^2 z}{\partial y^2} = b^2 k$$

3. If $z^3 + 3(ax + by)z = c^3$, prove that

$$x^2 \frac{\partial^2 z}{\partial x^2} + 2xy \frac{\partial^2 z}{\partial x \partial y} + y^2 \frac{\partial^2 z}{\partial y^2} = \frac{2z(ax + by)^3}{(ax + by + z^2)^3}$$

4. If $ax^3 + by^3 + cz^3 + 3hxyz = k$, show that

$$(hxy + cz^2)^3 \frac{\partial^2 z}{\partial x \partial y} = hk(hxy - cz^2) - 2(abc + h^3)x^2y^2z.$$

5. Given that $u = \dfrac{1}{\sqrt{x^2 + y^2 + z^2}}$, verify that

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 z}{\partial z^2} = 0$$

6. If $z = a \tan^{-1}(y/x)$, show that

   (i) $(1 + q^2)r - 2pqs + (1 + p^2)t = 0$

   (ii) $(rt - s^2)/(1 + p^2 + q^2)^2 = -a^2/(x^2 + y^2 + a^2)^2$

   where p, q, r, s, t have their usual meaning as in Example 23.

7. If $u = y^2 + z^{-2}$, $v = z^2 + x^{-2}$, $w = x^2 + y^{-2}$ and if V is a function of x, y, z; prove that

$$x\frac{\partial V}{\partial x} + y\frac{\partial V}{\partial y} + z\frac{\partial V}{\partial z} + 2\left(u\frac{\partial V}{\partial u} + v\frac{\partial V}{\partial v} + w\frac{\partial V}{\partial w}\right) = 4\left(y^2\frac{\partial V}{\partial u} + z^2\frac{\partial V}{\partial v} + x^2\frac{\partial V}{\partial w}\right)$$

8. If $z = f[(ny - mz)/(nx - lz)]$, prove that

$$(nx - lz)\frac{\partial z}{\partial x} + (ny - mz)\frac{\partial z}{\partial y} = 0$$

9. If $u = f(x + 2y) + g(x - 2y)$, show that

$$4\frac{\partial^2 u}{\partial x^2} = \frac{\partial^2 u}{\partial y^2}$$

10. If $u = \phi(x + at) + \psi(x - at)$, show that

$$\frac{\partial^2 u}{\partial t^2} = a^2\frac{\partial^2 u}{\partial x^2}.$$

    Prove that if $y = x + at$, $z = x - at$, the equation becomes $\dfrac{\partial^2 u}{\partial y \partial z} = 0$.

11. Given that $u = F(x, y, z)$ and $z = f(x, y)$, find $\dfrac{\partial^2 u}{\partial y^2}$ and $\dfrac{\partial^2 u}{\partial y \partial x}$ in terms of the derivatives of F and f (as in Example 24).

12. If $V = F(x, y)$ and $x = e^u \cos t$, $y = e^u \sin t$, show that

$$\frac{\partial^2 V}{\partial u^2} + \frac{\partial^2 V}{\partial t^2} = e^{2u}\left(\frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2}\right).$$

13. If $V = F(x, y)$ and $x = \frac{1}{2}u(e^v + e^{-v})$, $y = \frac{1}{2}u(e^v - e^{-v})$, show that

$$V_{xx} - V_{yy} = V_{uu} + \frac{1}{u}V_u + \frac{1}{u^2}V_{vv}.$$

14. If z is a function of u and v, and $u = x^2 - y^2 - 2xy$, $v = y$, prove that the equation

$$(x+y)\frac{\partial z}{\partial x} + (x-y)\frac{\partial z}{\partial y} = 0 \text{ is equivalent to } \frac{\partial z}{\partial v} = 0.$$

15. If $x = u + v$, $y = uv$ and z is a function of x and y, prove that

$$\frac{\partial^2 z}{\partial u^2} - 2\frac{\partial^2 z}{\partial u \partial v} + \frac{\partial^2 z}{\partial v^2} = (x^2 - 4y)\frac{\partial^2 z}{\partial y^2} - 2\frac{\partial z}{\partial y}$$

16. If $x = r\cos\theta$, $y = r\sin\theta$, prove that the equation

$$xy\left(\frac{\partial^2 u}{\partial x^2} - \frac{\partial^2 u}{\partial y^2}\right) - (x^2 - y^2)\frac{\partial^2 u}{\partial x \partial y} = 0$$

becomes

$$r\frac{\partial^2 u}{\partial r \partial \theta} - \frac{\partial u}{\partial \theta} = 0.$$

17. If $x = c\cosh u \cos v$, $y = c \sinh u \sin v$, and F is a function of x and y, show that

$$F_{uu} + F_{vv} = \frac{1}{2}c^2(\cosh 2u - \cos 2v)(F_{xx} + F_{yy}).$$

18. If V is a function of u, v, and $u = x^2 - y^2$, $v = 2xy$, prove that

$$4(u^2 + v^2)\frac{\partial^2 V}{\partial u \partial v} + 2u\frac{\partial V}{\partial v} + 2v\frac{\partial V}{\partial u} = xy\left(\frac{\partial^2 V}{\partial x \partial y} - \frac{\partial^2 V}{\partial y^2}\right) + \frac{1}{2}(x^2 - y^2)\frac{\partial^2 V}{\partial x \partial y}$$

19. Given that f is a function of x and y and that $x = u^2 v$, $y = uv^2$, prove that

$$2x^2 f_{x^2} + 2y^2 f_{y^2} + 5xyf_{xy} = uvf_{uv} - \frac{2}{3}(uf_u + vf_v)$$

20. Prove that, if in the equation

$$\frac{\partial^2 z}{\partial x^2} + 2xy^2\frac{\partial z}{\partial x} + 2(y - y^3)\frac{\partial z}{\partial y} + x^2y^2z^2 = 0,$$

the variables x, y are changed to u, v, where $x = uv$, $y = \frac{1}{v}$, the new equation is obtained by writing u for x and v for y, then z is the same function of u, v as of x, y.

21. If the variables x, y in the equations

$$(x^2 + y)^2\left(\frac{\partial^2 z}{\partial x^2} + \frac{\partial^2 z}{\partial y^2}\right) + 4xy\frac{\partial^2 z}{\partial x \partial y} + 2x\frac{\partial z}{\partial x} + 2y\frac{\partial z}{\partial y} = 0$$

are changed to u, v, where $2x = e^u + e^v$, $2y = e^u - e^v$, show that the new equation is

$$\frac{\partial^2 z}{\partial u^2} + \frac{\partial^2 z}{\partial v^2} = 0$$

22. If $x = r\sin\theta\cos\phi$, $y = r\sin\theta\sin\phi$, $z = r\cos\theta$, prove that

$$\frac{\partial r}{\partial x} = \sin\theta\cos\phi, \frac{\partial r}{\partial y} = \sin\theta\sin\phi, \frac{\partial r}{\partial z} = \cos\theta,$$

$$\frac{\partial \theta}{\partial x} = \frac{1}{r}\cos\theta\cos\phi,\ \frac{\partial \theta}{\partial y} = \frac{1}{r}\cos\theta\sin\phi,\ \frac{\partial \theta}{\partial z} = -\frac{1}{r}\sin\theta,$$

$$\frac{\partial \phi}{\partial x} = -\frac{\sin\phi}{r\sin\theta},\ \frac{\partial \phi}{\partial y} = \frac{\cos\phi}{r\sin\theta},\ \frac{\partial \phi}{\partial z} = 0.$$

Find also the derivatives of x, y, z with respect to $r,\ \theta,\ \phi$.

23. If u is a function of x, y, z, prove using values in Ex. 22, that

$$\left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2 + \left(\frac{\partial u}{\partial z}\right)^2 = \left(\frac{\partial u}{\partial r}\right)^2 + \left(\frac{1}{r}\frac{\partial u}{\partial \theta}\right)^2 + \left(\frac{1}{r\sin\theta}\frac{\partial u}{\partial \phi}\right)^2.$$

9. TAYLOR'S THEOREM

If $f(x, y)$ *is a function which possesses continuous partial derivatives of order n in any domain of a point* (a, b), *and the domain is large enough to contain a point* $(a+h, b+k)$ *with it, then there exists a positive number* $0 < \theta < 1$, *such that*

$$f(a+h, b+k) = f(a, b) + \left(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right) f(a, b) + \frac{1}{2!}\left(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right)^2 f(a, b)$$

$$+ \ldots + \frac{1}{(n-1)!}\left(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right)^{n-1} f(a, b) + R_n,$$

where $R_n = \dfrac{1}{n!}\left(h\dfrac{\partial}{\partial x} + k\dfrac{\partial}{\partial y}\right)^n f(a+\theta h, b+\theta k),\ 0 < \theta < 1.$

Let $x = a + th,\ y = b + tk$, where $0 \le t \le 1$ is a parameter, and

$$f(x, y) = f(a+th, b+tk) = \phi(t)$$

Since the partial derivatives of $f(x, y)$ of order n are continuous in the domain under consideration, $\phi^n(t)$ is continuous in [0, 1], and also

$$\phi'(t) = \frac{df}{dt} = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt} = h\frac{\partial f}{\partial x} + k\frac{\partial f}{\partial y} = \left(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right) f$$

$$\phi''(t) = \left(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right)^2 f$$

$$\vdots$$

$$\phi^{(n)}(t) = \left(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right)^n f$$

therefore by Maclaurin's theorem

$$\phi(t) = \phi(0) + t\phi'(0) + \frac{t^2}{2!}\phi''(0) + \dots + \frac{t^{n-1}}{(n-1)!}\phi^{(n-1)}(0) + \frac{t^n}{n!}\phi^{(n)}(\theta t),$$

where $0 < \theta < 1$.

Now on putting $t = 1$, we get

$$\phi(1) = \phi(0) + \phi'(0) + \frac{1}{2!}\phi''(0) + \dots + \frac{1}{(n-1)!}\phi^{(n-1)}(0) + \frac{1}{n!}\phi^{(n)}(0)$$

But $\phi(1) = f(a+h, b+k)$, and $\phi(0) = f(a, b)$

$$\phi'(0) = \left(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right)f(a, b)$$

$$\phi''(0) = \left(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right)^2 f(a, b)$$

$$\vdots$$

$$\phi^{(n)}(\theta) = \left(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right)^n f(a + \theta h, b + \theta k)$$

$$\therefore \quad f(a+h, b+k) = f(a, b) + \left(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right)f(a, b)$$

$$+ \frac{1}{2!}\left(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right)^2 f(a, b) + \dots + \frac{1}{(n-1)!}\left(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right)^{n-1} f(a, b) + R_n,$$

where $R_n = \dfrac{1}{n!}\left(h\dfrac{\partial}{\partial x} + k\dfrac{\partial}{\partial y}\right)^n f(a + \theta h, b + \theta k),\ 0 < \theta < 1.$

R_n is called the *remainder after n terms*, and the theorem, *Taylor's theorem with remainder* or *Taylor's expansion* about the point (a, b).

If we put $a = b = 0$; $h = x$, $k = y$, we get

$$f(x, y) = f(0, 0) + \left(x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y}\right)f(0, 0)$$

$$+ \frac{1}{2!}\left(x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y}\right)^2 f(0, 0) + \dots + \frac{1}{(n-1)!}\left(x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y}\right)^{n-1} f(0, 0) + R_n$$

where $R_n = \dfrac{1}{n!}\left(x\dfrac{\partial}{\partial x} + y\dfrac{\partial}{\partial y}\right)^n f(\theta x, \theta y),\ 0 < \theta < 1,$ is called the *Maclaurin's theorem* or *Maclaurin's expansion*.

It is easy to see that Taylor's theorem can also be put in the form:

$$f(a+h, b+k) = f(a, b) + df(a, b) + \frac{1}{2!}d^2 f(a, b) + \dots$$
$$+ \frac{1}{(n-1)!}d^{n-1} f(a, b) + \frac{1}{n!}d^n f(a + \theta h, b + \theta k)$$

The reasoning in the general case of several variables is precisely the same and so the theorem can be easily extended to any number of variables.

9.1 The Theorem can be Stated in Still another Form

$$f(x, y) = f(a, b) + \left[(x-a)\frac{\partial}{\partial x} + (y-b)\frac{\partial}{\partial y}\right] f(a, b)$$
$$+ \frac{1}{2!}\left[(x-a)\frac{\partial}{\partial x} + (y-b)\frac{\partial}{\partial y}\right]^2 f(a, b) + \dots$$
$$+ \frac{1}{(n-1)!}\left[(x-a)\frac{\partial}{\partial x} + (y-b)\frac{\partial}{\partial y}\right]^{n-1} f(a, b) + R_n,$$

where $R_n = \dfrac{1}{n!}\left[(x-a)\dfrac{\partial}{\partial x} + (y-b)\dfrac{\partial}{\partial y}\right]^n f(a + (x-a)\theta, b + (y-b)\theta),\ 0 < \theta < 1,$ called the Taylor's expansion of $f(x, y)$ about the point (a, b) in powers of $x - a$ and $y - b$.

Example 30. Expand $x^2y + 3y - 2$ in powers of $x - 1$ and $y + 2$.

■ Let us use Taylor's expansion with $a = 1$, $b = -2$. Then

$$f(x, y) = x^2y + 3y - 2, \qquad f(1, -2) = -10$$
$$f_x(x, y) = 2xy, \qquad f_x(1, -2) = -4$$
$$f_y(x, y) = x^2 + 3, \qquad f_y(1, -2) = 4$$
$$f_{xx}(x, y) = 2y, \qquad f_{xx}(1, -2) = -4$$
$$f_{xy}(x, y) = 2x, \qquad f_{xy}(1, -2) = 2$$
$$f_{yy}(x, y) = 0, \qquad f_{yy}(1, -2) = 0$$
$$f_{xxx}(x, y) = 0 = f_{yyy}(x, y), \qquad f_{yxx}(1, -2) = 2 = f_{xxy}(1, -2)$$

CHAPTER 15

All higher derivatives are zero.

$$\therefore \quad x^2y + 3y - 2 = -10 - 4(x-1) + 4(y+2) + \frac{1}{2}[-4(x-1)^2 + 4(x-1)(y+2)]$$

$$+ \frac{1}{3!}3(x-1)^2(y+2)(2) + 0$$

$$= -10 - 4(x-1) + 4(y+2) - 2(x-1)^2 + 2(x-1)(y+2) + (x-1)^2(y+2)$$

Example 31. If $f(x, y) = \sqrt{|xy|}$, prove that Taylor's expansion about the point (x, x) is not valid in any domain which includes the origin.

■ As was shown earlier in Example II § 4.1,

$$f_x(0, 0) = 0 = f_y(0, 0)$$

$$f_x(x, y) = \begin{cases} \frac{1}{2}\sqrt{|y/x|}, & x > 0 \\ -\frac{1}{2}\sqrt{|y/x|}, & x < 0 \end{cases}$$

$$f_y(x, y) = \begin{cases} \frac{1}{2}\sqrt{|x/y|}, & y > 0 \\ -\frac{1}{2}\sqrt{|x/y|}, & y < 0 \end{cases}$$

$$\therefore \quad f_x(x, x) = f_y(x, x) = \begin{cases} \frac{1}{2}, & x > 0 \\ -\frac{1}{2}, & x < 0 \end{cases}$$

Now Taylor's expansion about (x, x) for $n = 1$, is

$$f(x+h, x+h) = f(x, x) + h[f_x(x+\theta h, x+\theta h) + f_y(x+\theta h, x+\theta h)]$$

or

$$|x+h| = \begin{cases} |x| + h, & \text{if } x + \theta h > 0 \\ |x| - h, & \text{if } x + \theta h < 0 \\ |x|, & \text{if } x + \theta h = 0 \end{cases} \qquad ...(1)$$

If the domain $(x, x; x+h, x+h)$ includes the origin, then x and $x+h$ must be of opposite signs, that is either

$$|x+h| = x + h, \quad |x| = -x$$

or

$$|x+h| = -(x+h), \quad |x| = x$$

But under these conditions none of the inequalities (1) holds. Hence the expansion is not valid.

Ex. 1. Expand $x^4 + x^2y^2 - y^4$ about the point $(1, 1)$ up to terms of the second degree. Find the form of R_2.

Ex. 2. Find the expansion of $\sin x \sin y$ about $(0, 0)$ up to and including the terms of the fourth degree in (x, y). Compare the result with that you get by multiplying the series for $\sin x$ and $\sin y$.

Ex. 3. Expand $e^x \tan^{-1} y$ about $(1, 1)$ up to the second degree in $(x-1)$ and $(y-1)$.

Ex. 4. Show that the expansion of $\sin(xy)$ in powers of $(x-1)$ and $(y-\pi/2)$ up to and including second degree terms is

$$1-\frac{1}{8}\pi^2(x-1)^2-\frac{1}{2}\pi(x-1)\left(y-\frac{\pi}{2}\right)-\frac{1}{2}\left(y-\frac{\pi}{2}\right)^2$$

Ex. 5. Show that, for $0<\theta<1$,

$$\sin x \sin y = xy-\frac{1}{6}[(x^3+3xy^2)\cos\theta x\sin\theta y+(y^3+3x^2y)\sin\theta x\cos\theta y]$$

Ex. 6. Prove that the first four terms of the Maclaurin expansion of $e^{ax}\cos by$ are

$$1+ax+\frac{a^2x^2-b^2y^2}{2!}+\frac{a^3x^3-3ab^2xy^2}{3!}.$$

Ex. 7. Prove that for $0<\theta<1$,

$$e^{ax}\sin by = by+abxy+\frac{1}{6}[(a^3x^3-3ab^2xy^2)\sin(b\theta y)$$
$$+(3a^2bx^2y-b^3y^3)\cos(b\theta y)]\,e^{a\theta x}$$

Ex. 8. Show that if f, f_x, f_y are all continuous in a domain D of (a, b), and D is large enough to contain the point $(a+h, b+k)$, within it, then for $0<\theta<1$,

$$f(a+h, b+k) = f(a, b)+hf_x(a+\theta h, b+\theta k)+kf_y(a+\theta h, b+\theta k).$$

If $f(x, y) = x\sqrt{x^2+y^2}$, $a=b=-1$, $h=k=3$, verify that the above conditions are satisfied and find the value of θ.

10. EXTREME VALUES: MAXIMA AND MINIMA

The theory of extreme values (maximum or minimum) for functions of one variable was considered in an earlier chapter. We now investigate the theory for explicit functions of more than one variable. That for implicit functions will be discussed in the next chapter.

Let (a, b) be a point in the domain of definition of a function f. Then $f(a, b)$ is an *extreme value* of f, if for every point (x, y), [other than (a, b)] of some neighbourhood of (a, b), the difference

$$f(x, y)-f(a, b) \qquad \text{...(1)}$$

keeps the same sign.

The extreme value $f(a, b)$ is called *a maximum or a minimum value* according as the sign of (1) is negative or positive.

10.1 A Necessary Condition

A necessary condition for $f(x, y)$ to have an extreme value at (a, b) is that $f_x(a, b) = 0, f_y(a, b) = 0$, provided these partial derivatives exist.

If $f(a, b)$ is an extreme value of the function $f(x, y)$ of two variables, then it must also be an extreme value of both the functions, $f(x, b)$ and $f(a, y)$ of one variable. But a necessary condition that these have extreme value at $x = a$ and $y = b$ respectively, is

$$f_x(a, b) = 0,\ f_y(a, b) = 0.$$

Notes:

1. The function $f(x, y) = |x| + |y|$ has an extreme value at $(0, 0)$ even though the partial derivatives f_x and f_y do not exist at $(0, 0)$.
2. If $f(x, y) = 0$, if $x = 0$ or $y = 0$, and $f(x, y) = 1$ elsewhere, then both the partial derivatives exist (each equal to zero) at the origin, but $f(0, 0)$ is not an extreme value. Thus the conditions obtained above are *only necessary and not sufficient*.
3. Point at which $f_x = 0, f_y = 0$ (or $df = 0$) are called *Stationary points*.

10.2 Sufficient Conditions for $f(x, y)$ to have an Extreme Value at (a, b)

Let $f_x(a, b) = 0 = f_y(a, b)$. Further, let us suppose that $f(x, y)$ possesses continuous second order partial derivatives in a certain neighbourhood of (a, b) and that these derivatives at (a, b) viz. $f_{xx}(a, b)$, $f_{xy}(a, b), f_{yy}(a, b)$ are not all zero.

Let $(a + h, b + k)$ be a point of this neighbourhood.

Let us write

$$A = f_{xx}(a, b),\ B = f_{xy}(a, b),\ C = f_{yy}(a, b)$$

By Taylor's theorem, we have for $0 < \theta < 1$,

$$f(a + h, b + k) = f(a, b) + [hf_x(a, b) + kf_y(a, b)]$$

$$+ \frac{1}{2!}[h^2 f_{xx}(a + \theta h, b + \theta k) + 2hkf_{xy}(a + \theta h, b + \theta k) + k^2 f_{yy}(a + \theta h, + b + \theta k)]$$

But $f_x(a, b) = 0 = f_y(a, b)$, and

Since the second order partial derivatives are continuous at (a, b), we write

$$f_{xx}(a + \theta h, b + \theta k) - f_{xx}(a, b) = \rho_1$$

$$f_{xy}(a + \theta h, b + \theta k) - f_{xy}(a, b) = \rho_2$$

$$f_{yy}(a + \theta h, b + \theta k) - f_{yy}(a, b) = \rho_3$$

where ρ_1, ρ_2, ρ_3 are functions of h and k, and $\to 0$ as $(h, k) \to (0, 0)$.

$$\therefore \quad f(a + h, b + k) - f(a, b) = \frac{1}{2}[Ah^2 + 2Bhk + Ck^2 + \rho]$$

where $\rho = \rho_1 + \rho_2 + \rho_3 \to 0$ as $(h, k) \to (0, 0)$ and is of unknown sign.

Let $G = Ah^2 + 2Bhk + Ck^2$, so that

$$f(a+h, b+k) - f(a, b) = \frac{1}{2}[G + \rho] \qquad \text{...(1)}$$

There are several cases to consider.

(*i*) G never vanishes and keeps a constant sign. Since $\rho \to 0$ when $(h, k) \to (0, 0)$, therefore ρ is a small number and the sign of $G + \rho$ is same as of G, *i.e.*, $G + \rho$ remains negative or positive according as G is negative or positive. Thus the difference,

$f(a+h, b+k) - f(a, b) \lesseqgtr 0$ according as $G \lesseqgtr 0$

But we know by definition that $f(x, y)$ has a maxima or a minima at (a, b) according as the difference $f(a+h, b+k) - f(a, b)$ is negative or positive for all (h, k) except $(0, 0)$.

Thus $f(a, b)$ will be a maximum or a minimum value according as G is negative or positive.

(*ii*) If G can change sign, since $f(a+h, b+k) - f(a, b)$ and G have the same sign when ρ is small, $f(a, b)$ will not be an extreme value.

(*iii*) If G, without ever changing sign, vanishes for certain values of (h, k), the sign of $f(a+h, b+k) - f(a, b)$ will depend upon ρ, which is of unknown sign, and so no conclusion can be drawn. This is the *doubtful case* and requires further investigation.

Let us first take $A \neq 0$.

G may be written in the form:

$$G = \frac{(Ah + Bk)^2 + k^2(AC - B^2)}{A}$$

(1) If $AC - B^2 > 0$, the numerator of G is the sum of two positive quantities and it never vanishes except when $k = 0$, $h = 0$, simultaneously, which is not permissible [see(*i*)]. Hence, G never vanishes and has the same sign as A.

Thus, $f(a, b)$ has a maximum value if $A < 0$, and a minimum value if $A > 0$.

(2) If $AC - B^2 < 0$, the sign of the numerator of G may be positive or negative according as $(Ah + Bk)^2 >$ or $< k^2(B^2 - AC)$, *i.e.*, according to the values of (h, k). Hence, G does not keep the same sign for all values of (h, k), and therefore, $f(a, b)$ is not an extreme value.

(3) If $AC - B^2 = 0$, the numerator of G is a perfect square but may vanish for values of (h, k) for which $Ah + Bk = 0$. Thus G, without changing sign, may vanish for certain values of (h, k).

This is the doubtful case in which the sign of $f(a+h, b+k) - f(a, b)$ depends upon ρ and requires further investigation.

If $A = 0$, then

$$G = 2Bhk + Ck^2 = k(2Bh + Ck)$$

(4) If $A = 0$, $B \neq 0$, G changes sign with k and $(2Bh + Ck)$, and there is no extreme value.

(5) If $A = 0$, $B = 0$, G does not change sign but may vanish when $k = 0$ (without $h = 0$). This is therefore the doubtful case and requires further investigation.

Rule. *$f(a, b)$ is an extreme value of $f(x, y)$, if $f_x(a, b) = 0 = f_y(a, b)$, and*

$$f_{xx}(a, b) \cdot f_{yy}(a, b) - [f_{xy}(a, b)]^2 > 0,$$

and this extreme value is a maximum or a minimum according as $f_{xx}(a, b)$ [or $f_{yy}(a, b)$] is negative or positive.

Further investigation is necessary, if

$$f_{xx}(a, b) f_{yy}(a, b) - [f_{xy}(a, b)]^2 = 0$$

Remark: Since at (a, b),

$$df = hf_x(a, b) + kf_y(a, b)$$

and

$$d^2f = h^2 f_{xx}(a, b) + 2hkf_{xy}(a, b) + k^2 f_{yy}(a, b) = Ah^2 + 2Bhk + Ck^2$$

so $f(a, b)$ is an extreme value of $f(x, y)$ if at (a, b), $df = 0$, and d^2f keeps the same sign for all values of $(h, k) \neq (0, 0)$.

Note: Discussion of the doubtful case involves the consideration of terms of higher order than the second in the Taylor expansion of $f(a + h, b + k)$ but this is generally not easy and will not be considered here.

However, it is sometimes possible to decide whether f has a maxima or a minima at (a, b) by algebraic or geometric considerations.

Example 32. Find the maxima and minima of the function

$$f(x, y) = x^3 + y^3 - 3x - 12y + 20$$

■ We have

$$f_x(x, y) = 3x^2 - 3 = 0, \text{ when } x = \pm 1$$

$$f_y(x, y) = 3y^2 - 12 = 0, \text{ when } y = \pm 2$$

Thus, the function has four stationary points:

$$(1, 2), (-1, 2), (1, -2), (-1, -2)$$

Now

$$f_{xx}(x, y) = 6x, f_{xy}(x, y) = 0, f_{yy}(x, y) = 6y$$

At $(1, 2)$

$$f_{xx} = 6 > 0, \text{ and } f_{xx} f_{yy} - (f_{xy})^2 = 72 > 0$$

Hence, $(1, 2)$ is a point of minima of the function.

At $(-1, 2)$

$$f_{xx} = -6, \text{ and } f_{xx} f_{yy} - (f_{xy})^2 = -72 < 0$$

Hence, the function has neither a maxima nor a minima at $(-1, 2)$.

At $(1, -2)$,

$$f_{xx} = 6, \text{ and } f_{xx} f_{yy} - (f_{xy})^2 = -72 < 0.$$

Hence, the function has neither maximum nor minimum at $(1, -2)$.

At (–1, –2),

$$f_{xx} = -6, \text{ and } f_{xx}f_{yy} - (f_{xy})^2 = 72 > 0$$

Hence, the function has a maximum value at (–1, –2).

Note: Stationary points like (–1, 2), (1, –2) which are not extreme points are called the *saddle points*.

Example 33. Show that the function

$$f(x, y) = 2x^4 - 3x^2y + y^2$$

has neither a maximum nor a minimum at (0, 0), where

$$f_{xx}f_{yy} - (f_{xy})^2 = 0.$$

■ Now

$$f_x(x, y) = 8x^3 - 6xy, \; f_y(x, y) = -3x^2 + 2y$$

$\therefore$
$$f_x(0, 0) = 0 = f_y(0, 0)$$

Also

$$f_{xx}(x, y) = 24x^2 - 6y = 0, \text{ at } (0, 0)$$
$$f_{xy}(x, y) = -6x = 0, \text{ at } (0, 0)$$
$$f_{yy}(x, y) = 2, \text{ at } (0, 0)$$

Thus at (0, 0), $f_{xx}(0, 0) \cdot f_{yy}(0, 0) - [f_{xy}(0, 0)]^2 = 0$.

So that it is a *doubtful* case, and thus requires further examination.

Again

$$f(x, y) = (x^2 - y)(2x^2 - y); \; f(0, 0) = 0$$

or

$$\begin{aligned} f(x, y) - f(0, 0) &= (x^2 - y)(2x^2 - y) \\ &> 0, \text{ for } y < 0 \text{ or } x^2 > y > 0 \\ &< 0, \text{ for } y > x^2 > \frac{y}{2} > 0 \end{aligned}$$

Thus $f(x, y) - f(0, 0)$ does not keep the same sign near the origin. Hence f has neither a maximum nor a minimum value at the origin.

Example 34. Show that

$$f(x, y) = y^2 + x^2y + x^4, \text{ has a minimum at } (0, 0).$$

■ It can be easily verified that at the origin,

$$f_x = 0, f_y = 0, f_{xx} = 0, f_{xy} = 0, f_{yy} = 2.$$

Thus at the origin $f_{xx}f_{yy} - (f_{xy})^2 = 0$, so that it is a *doubtful* case and requires further investigation.

But we can write

$$f(x, y) = \left(y + \frac{1}{2}x^2\right)^2 + \frac{3}{4}x^4$$

and

$$f(x, y) - f(0, 0) = \left(y + \frac{1}{2}x^2\right)^2 + \frac{3}{4}x^4$$

which is greater than zero for all values of (x, y). Hence f has a minimum value at the origin.

EXERCISE

1. Examine the following for extreme values:

   (*i*) $4x^2 - xy + 4y^2 + x^3y + xy^3 - 4$ (*ii*) $x^3y^2(12 - 3x - 4y)$

   (*iii*) $y^2 + 4xy + 3x^2 + x^3$ (*iv*) $(x^2 + y^2 - 4)^2 - x^2$

   (*v*) $(x^2 + y^2)\, e^{6x+2x^2}$ (*vi*) $(x - y)^2 (x^2 + y^2 - 2)$

   (*vii*) $x^3 + y^3 - 63(x + y) + 12xy.1$

2. Investigate the maxima and minima of the functions,

   (*i*) $21x - 12x^2 - 2y^2 + x^3 + xy^2$ (*ii*) $2(x - y)^2 - x^4 - y^4$

   (*iii*) $x^2 + 3xy + y^2 + x^3 + y^3$ (*iv*) $x^2 + 4xy + 4y^2 + x^3 + 2x^2y + y^4$

   (*v*) $x^2y^2 - 5x^2 - 8xy - 5y^2$ (*vi*) $x^2 - 2xy + y^2 + x^3 - y^3 + x^5$

3. Show that the function $(y - x)^4 + (x - 2)^4$ has a minimum at (2, 2).

4. Prove that the function $f(x, y) = x^2 - 2xy + y^2 + x^4 + y^4$ has a minima at the origin.

5. Find and classify the extreme values (if any) of the functions:

   (*i*) $x^2 + y^2 + x + y + xy$ (*ii*) $x^2 + xy + y^2 + ax + by$

   (*iii*) $y^2 - x^3$ (*iv*) $x^4 + y^4 - 6(x^2 + y^2) + 8xy$

6. A rectangular box, open at the top, is to have a volume of 32 cu ft. What must be the dimensions so that the total surface is a minimum.

7. Show that the function $f(x, y) = x^2 - 3xy^2 + 2y^4$ has neither a maximum nor a minimum value at the origin.

   [*Hint*: $f(x, y) = (x - y^2)(x - 2y^2)$].

ANSWERS

1. (*i*) min. at (0, 0), max. at $(\pm 3/2, \mp 3/2)$; (*ii*) max. at (2, 1);

   (*iii*) min. at (2/3, –4/3); (*iv*) min. at $\left(\pm 3\sqrt{2}/2, 0\right)$;

   (*v*) extremes at (0, 0), (–1, 0), (–1/2, 0); (*vi*) min. at $\left(\mp 1/\sqrt{2}, \pm 1/\sqrt{2}\right)$;

   (*vii*) min. at (3, 3), max. at (–7, –7), and neither max. nor min. at (5, –1) and (–1, 5)

2. (*i*) min. at (7, 0), max. at (1, 0); (*ii*) max at (0, 0);

   (*iii*) max. at (–5/3, –5/3); (*iv*) min. at (0, 0);

   (*v*) max. at (0, 0); (*vi*) neither max. nor min. at (0, 0);

5. (*i*) min. at (–2/3, –2/3); (*ii*) min. at $\left[\frac{1}{3}(b-2a), \frac{1}{3}(a-2b)\right]$;

(*iii*) neither at (0, 0);

(*iv*) max. at (0, 0), min at $\left(\pm\sqrt{5}, \mp\sqrt{5}\right)$, neither at $(\pm 1, \mp 1)$.

11. FUNCTIONS OF SEVERAL VARIABLES

We conclude the chapter by referring briefly—in fact very briefly, to the functions of several variables.

An ordered set $(a_1, a_2, ..., a_n)$ of n real numbers is called a **point** in a space of n dimensions. The aggregate of points $(x_1, x_2, ..., x_n)$ when $x_1, x_2, ..., x_n$ range over the entire set of real numbers is referred to as the *space of n dimensions*, denoted by $\boldsymbol{R}^n$.

Neighbourhoods. The set of values $x_1, x_2, ..., x_n$ other than $a_1, a_2, ..., a_n$ that satisfy the conditions

$$|x_1 - a_1| < \delta, |x_2 - a_2| < \delta, ..., |x_n - a_n| < \delta$$

where δ is an arbitrarily small positive number, is said to form a *neighbourhood* of the point $(a_1, a_2, ..., a_n)$. The neighbourhood may however be specified in other, though equivalent ways.

The rectangle

$$(a_1 - h_1, a_1 + h_1; a_2 - h_2, a_2 + h_2; ..., a_n - h_n, a_n + h_n)$$

where $h_1, h_2, ..., h_n$ are arbitrarily small positive numbers, is said to be rectangular neighbourhood of $(a_1, a_2, ..., a_n)$.

The points inside the sphere $x_1^2 + x_2^2 + ... + x_n^2 = \delta^2$ may be taken as a neighbourhood of the point (0, 0, ..., 0), called a spherical *nbd.*

Continuity. A function $f(x_1, x_2, ..., x_n)$ is said to be *continuous at a point* $P(a_1, a_2, ..., a_n)$, if to every positive number ε, there corresponds a neighbourhood of P such that for every point $(x_1, x_2, ..., x_n)$ of this neighbourhood

$$|f(x_1, x_2, ..., x_n) - f(a_1, a_2, ..., a_n)| < \varepsilon$$

A function which is continuous at every point of a region is said to be *continuous in that region.*

Partial Derivatives

The derivative of a function with respect to one variable, when all others are kept constant, is called the *partial derivative* of the function with respect to that variable. Thus the partial derivative of f with respect to x_1 at $(a_1, a_2, ..., a_n)$ is

$$\lim_{h_1 \to 0} \frac{f(a_1 + h_1, a_2, ..., a_n) - f(a_1, a_2, ..., a_n)}{h_1}$$

and is denoted by $f_{x_1}(a_1, a_2, ..., a_n)$.

The other partial derivatives of the first, second or higher orders may be defined similarly.

Differentiability. A function f is said to be *differentiable* at $(a_1, a_2, ..., a_n)$, if the change δf in the value of the function, when the point changes from $(a_1, a_2, ..., a_n)$ to $(a_1 + h_1, a_2 + h_2, ..., a_n + h_n)$, can be expressed in the form

$$\delta f = A_1 h_1 + A_2 h_2 + ... + A_n h_n + h_1\phi_1 + h_2\phi_2 + ... + h_n\phi_n,$$

where $\phi_1, \phi_2, ..., \phi_n$ are functions of $h_1, h_2, ..., h_n$ and tend to zero as $(h_1, h_2, ..., h_n) \to (0, 0, ..., 0)$.

A function which is differentiable at every point of a region, is said to be differentiable over the region.

The differentials $df, d^2f, ...$, may now be easily defined as in the case of functions of two variables.

11.1 Extreme Values of a Function of *n* Variables

A point $(a_1, a_2, ..., a_n)$ is said to be an *extreme point*, and $f(a_1, a_2, ..., a_n)$ an *extreme value* of a function f, if for every point $(x_1, x_2, ..., x_n)$, other than $(a_1, a_2, ..., a_n)$ of some neighbourhood of $(a_1, a_2, ..., a_n)$, the difference,

$$f(x_1, x_2, ..., x_n) - f(a_1, a_2, ..., a_n) \qquad ...(1)$$

keeps the same sign. The extreme value is a **maximum** or a **minimum value** according as the sign is negative or positive.

If $f(a_1, a_2, ..., a_n)$ is an extreme value of the function f of n variables, then it must also be an extreme value of the function $f(x_1, a_2, ..., a_n)$ of one variable x_1 and therefore the partial derivative $f_{x_1}(a_1, a_2, ..., a_n)$, in case it exists, must be zero. The same is true for all the other variables $x_2, x_3, ..., x_n$.

Thus, the **necessary conditions** for $f(a_1, a_2, ..., a_n)$ to be an extreme value of the function f are that all the partial derivatives $f_{x_1}, f_{x_2}, ..., f_{x_n}$, in case they exist, vanish at $(a_1, a_2, ..., a_n)$.

Since these are only necessary and not sufficient conditions, therefore points which satisfy these conditions may not be extreme points. A point $(a_1, a_2, ..., a_n)$ is called a **stationary point** if all the first order partial derivatives of the function vanish at that point. Thus the stationary points are determined by solving the following n equations simultaneously.

$$\begin{aligned} &f_{x_1}(x_1, x_2, ..., x_n) = 0 \\ &f_{x_2}(x_1, x_2, ..., x_n) = 0 \\ &\vdots \\ &f_{x_n}(x_1, x_2, ..., x_n) = 0 \end{aligned}$$

For a function f of n independent variables $x_1, x_2, ..., x_n$ the condition can be given in a more compact form.

Thus, if $(a_1, a_2, ..., a_n)$ is a stationary point, then

$$df(a_1, a_2, ..., a_n) = 0$$

i.e., the differential of the function vanishes at a stationary point. For, at the stationary point all the partial derivatives vanish and therefore,

$$\begin{aligned} df(a_1, a_2, ..., a_n) &= f_{x_1}(a_1, a_2, ..., a_n)\,dx_1 + f_{x_2}(a_1, a_2, ..., a_n)\,dx_2 \\ &\quad + ... + f_{x_n}(a_1, a_2, ..., a_n)\,dx_n = 0 \end{aligned}$$

Conversely, when $df = 0$, the coefficients of the differentials $dx_1, dx_2, \ldots, dx_n$ of independent variables, are separately equal to zero.

Further investigations are necessary to decide whether a stationary point is an extreme point or not, or whether it is a maxima or a minima. We now state a rule (without giving a proof which is beyond the scope of this book) for a function of three variables but is applicable to a function of any number of variables.

The Rule

For a function $f(x, y, z)$ of three independent variables, sufficient conditions for (a, b, c) to be an extreme point are that

(*i*) $df(a, b, c) = f_x dx + f_y dy + f_z dz = 0$, *so that* $f_x = f_y = f_z = 0$, *and*

(*ii*) $d^2 f(a, b, c) = f_{xx}(dx)^2 + f_{yy}(dy)^2 + f_{zz}(dz)^2 + 2f_{xy} dx\,dy + 2f_{yz}\,dy\,dz + 2f_{zx} dz\,dx,$

keep the same sign for arbitrary values of dx, dy, dz; the extreme point being a maxima or a minima according as the sign of $d^2 f$ is negative or positive. The point will be neither a maxima nor a minima if $d^2 f$ does not keep the same sign; and requires further investigation, if $d^2 f$ keeps the same sign but vanishes at some points of a nbd of (a, b, c).

The conditions that $d^2 f$ keeps the same sign may be stated (without proof) in terms of matrices, as follows:

Consider the matrix

$$\begin{bmatrix} f_{xx} & f_{xy} & f_{xz} \\ f_{yx} & f_{yy} & f_{yz} \\ f_{zx} & f_{zy} & f_{zz} \end{bmatrix}$$

$d^2 f$ will always be positive if and only if the three principal minors

$$f_{xx}, \begin{vmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{vmatrix}, \begin{vmatrix} f_{xx} & f_{xy} & f_{xz} \\ f_{yx} & f_{yy} & f_{yz} \\ f_{zx} & f_{zy} & f_{zz} \end{vmatrix}$$

are all positive, and $d^2 f$ will be always negative if and only if their signs are alternatively negative and positive.

Example 35. Show that

$$f(x, y, z) = (x + y + z)^3 - 3(x + y + z) - 24xyz + a^3$$

has a minima at $(1, 1, 1)$ and a maxima at $(-1, -1, -1)$.

■ We have

$$f_x = 3(x + y + z)^2 - 24yz - 3$$
$$f_y = 3(x + y + z)^2 - 24zx - 3$$
$$f_z = 3(x + y + z)^2 - 24xy - 3$$

Hence, the stationary points are given by

$$\left.\begin{array}{l}(x+y+z)^2-8yz-1=0\\(x+y+z)^2-8zx-1=0\\(x+y+z)^2-8xy-1=0\end{array}\right\}$$

Subtracting second equation from the first,

$$z(x-y)=0.$$

Similarly, $x(y-z)=0$, $y(z-x)=0$.

Therefore, either $x=0$, $y=0$, $z=0$ or $x=y=z$

Hence, stationary points are $(1, 1, 1)$, and $(-1, -1, -1)$.

Again, we have

$$f_{xx}=6(x+y+z)=f_{yy}=f_{zz}$$
$$f_{xy}=6(x+y+z)-24z=f_{yx}$$
$$f_{yz}=6(x+y+z)-24x=f_{zy}$$
$$f_{zx}=6(x+y+z)-24x=f_{xz}$$

At $(1, 1, 1)$,

$$f_{xx}=f_{yy}=f_{zz}=18,\ f_{xy}=f_{yz}=f_{zx}=-6$$

$$\therefore \quad d^2f=18(dx^2+dy^2+dz^2)-12(dx\,dy+dy\,dz+dz\,dx)$$
$$=6[(dx^2+dy^2+dz^2)+(dx-dy)^2+(dy-dz)^2+(dz-dx)^2]$$

which is positive for all values of dx, dy, dz and does not vanish for

$$(dx, dy, dz)\neq(0,0,0)$$

Thus $(1, 1, 1)$ is a point of minima of the function.

At $(-1, -1, -1)$,

$$f_{xx}=f_{yy}=f_{zz}=-18,\ f_{xy}=f_{yz}=f_{zx}=6$$

$$\therefore \quad d^2f=-18(dx^2+dy^2+dz^2)+12(dx\,dy+dy\,dz+dz\,dx)$$
$$=-6[(dx^2+dy^2+dz^2)+(dx-dy)^2+(dy-dz)^2+(dz-dx)^2]$$

which is negative for all dx, dy, dz and never vanishes. Hence the function has a maximum value at $(-1, -1, -1)$.

Example 36. Show that the minimum and the maximum values of

$$f(x, y, z)=(ax+by+cz)\,e^{-\alpha^2x^2-\beta^2y^2-\gamma^2z^2} \text{ are}$$

$$-\sqrt{\frac{1}{2}(a^2\alpha^{-2}+b^2\beta^{-2}+c^2\gamma^{-2})/e} \text{ and } \sqrt{\frac{1}{2}(a^2\alpha^{-2}+b^2\beta^{-2}+c^2\gamma^{-2})/e}$$

■ We have

$$f_x=[a-2\alpha^2x\,\Sigma ax]\exp(-\Sigma\alpha^2x^2)$$
$$f_y=[b-2\beta^2y\,\Sigma ax]\exp(-\Sigma\alpha^2x^2)$$
$$f_z=[c-2\gamma^2z\,\Sigma ax]\exp(-\Sigma\alpha^2x^2)$$

At the stationary point, since $\exp(-\Sigma\alpha^2 x^2) \neq 0$, we have

$$\left.\begin{aligned} a - 2\alpha^2 x\,\Sigma ax = 0 \\ b - 2\beta^2 y\,\Sigma ax = 0 \\ c - 2\gamma^2 z\,\Sigma ax = 0 \end{aligned}\right\} \qquad \ldots(1)$$

$$\therefore \quad \left.\begin{aligned} x\Sigma ax = a/2\alpha^2 \\ x\Sigma ax = b/2\beta^2 \\ x\Sigma ax = c/2\gamma^2 \end{aligned}\right\}$$

Multiplying by a, b, c and adding,

$$(\Sigma ax)^2 = \frac{1}{2}\Sigma a^2\alpha^{-2}$$

$$\therefore \quad \Sigma ax = \sqrt{\frac{1}{2}\Sigma a^2\alpha^2} = \pm k, \text{ say}$$

Hence from (1), the stationary points are

$$\left(\frac{a}{2\alpha^2 k}, \frac{b}{2\beta^2 k}, \frac{c}{2\gamma^2 k}\right), \left(-\frac{a}{2\alpha^2 k}, -\frac{b}{2\beta^2 k}, -\frac{c}{2\gamma^2 k}\right)$$

Again, we have

$$f_{xx} = -2a^2 x[a - 2\alpha^2 x\Sigma ax]\exp(-\Sigma\alpha^2 x^2) - 2\alpha^2[\Sigma ax + ax]\exp(-\Sigma\alpha^2 x^2)$$

$$f_{xy} = -2a^2 x[b - 2\beta^2 y\Sigma ax]\exp(-\Sigma\alpha^2 x^2) - 2\beta^2 ay\exp(-\Sigma\alpha^2 x^2)$$

and similar expressions for $f_{yy}, f_{zz}, f_{yz}, f_{zx}$,

At the stationary point $\left(\frac{a}{2\alpha^2 k}, \frac{b}{2\beta^2 k}, \frac{c}{2\gamma^2 k}\right)$, we have $\Sigma\alpha^2 x^2 = \frac{1}{2}$,

$$f_{xx} = 0 - 2\alpha^2\left[k + \frac{a^2}{2\alpha^2 k}\right]e^{-1/2} = -\frac{1}{\sqrt{e}}\left[2\alpha^2 k + \frac{a^2}{k}\right] = -\frac{2\alpha^2 k^2 + a^2}{k\sqrt{e}}$$

$$f_{yy} = -\frac{2\beta^2 k^2 + b^2}{k\sqrt{e}}, \quad f_{zz} = -\frac{2\gamma^2 k^2 + c^2}{k\sqrt{e}}$$

$$f_{xy} = 0 - \frac{ab}{k\sqrt{e}}, \quad f_{yz} = -\frac{bc}{k\sqrt{e}}, \quad f_{zx} = -\frac{ca}{k\sqrt{e}}$$

$$\therefore \quad d^2 f = \frac{-1}{k\sqrt{e}}[(2\alpha^2 k^2 + a^2)dx^2 + (2\beta^2 k^2 + b^2)dy^2 + (2\gamma^2 k^2 + c^2)dz^2]$$

$$-\frac{2}{k\sqrt{e}}[ab\,dxdy + bc\,dydz + ca\,dzdx]$$

Now $f_{xx} < 0$

$$\begin{vmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{vmatrix} = \begin{vmatrix} -\dfrac{2\alpha^2k^2+\alpha^2}{k\sqrt{e}} & -\dfrac{ab}{k\sqrt{e}} \\ -\dfrac{ab}{k\sqrt{e}} & -\dfrac{2\beta^2k^2+b^2}{k\sqrt{e}} \end{vmatrix} = \frac{1}{k^2e}\begin{vmatrix} 2\alpha^2k^2+a^2 & ab \\ ab & 2\beta^2k^2+b^2 \end{vmatrix}$$

$$= \frac{2}{e}(2\alpha^2\beta^2k^2 + \alpha^2b^2 + a^2\beta^2) > 0$$

$$\begin{vmatrix} f_{xx} & f_{xy} & f_{xz} \\ f_{yx} & f_{yy} & f_{yz} \\ f_{zx} & f_{zy} & f_{zz} \end{vmatrix} = -\frac{1}{k^3e^{3/2}}\begin{vmatrix} 2\alpha^2k^2+a^2 & ab & ac \\ ab & 2\beta^2k^2+b^2 & bc \\ ac & bc & 2\gamma^2k^2+c^2 \end{vmatrix}$$

$$= -\frac{1}{k^3e^{3/2}}[(2\gamma^2k^2+c^2)2k^2(2\alpha^2\beta^2k^2+\alpha^2b^2+a^2\beta^2) + a^2c^2(-2\beta^2k^2) - 2\alpha^2b^2c^2k^2]$$

$$= -\frac{4k}{e^{3/2}}[2\alpha^2\beta^2\gamma^2k^2 + \alpha^2b^2c^2 + \alpha^2\beta^2c^2 + a^2b^2\gamma^2] < 0.$$

Thus, the three principal minors have alternatively negative and positive signs and so d^2f is always negative, and hence $(a/2\alpha^2k, b/2\beta^2k, c/2\gamma^2k)$ is a point of maxima, and the maximum value $= ke^{-1/2} = \sqrt{\frac{1}{2}\Sigma a^2\alpha^{-2}/e}$.

At the point $(-a/2\alpha^2k, -b/2\beta^2k, -c/2\gamma^2k)$, it may be shown as above that $\Sigma\alpha^2x^2 = \frac{1}{2}$, and

$$f_{xx} = \frac{2\alpha^2k^2+a^2}{k\sqrt{e}},\ f_{yy} = \frac{2\beta^2k^2+b^2}{k\sqrt{e}},\ f_{zz} = \frac{2\gamma^2k^2+c^2}{k\sqrt{e}}$$

$$f_{xy} = \frac{ab}{k\sqrt{e}},\ f_{yz} = \frac{bc}{k\sqrt{e}},\ f_{zx} = \frac{ca}{k\sqrt{e}}$$

and the three principal minors are of positive sign, so that d^2f is positive, and hence the point in question is a minima and the minimum value of the function $= -ke^{-1/2} = -\sqrt{\frac{1}{2}\Sigma a^2\alpha^{-2}/e}$.

EXERCISE

1. Show that the following functions have a minima at the points indicated:

   (i) $x^2 + y^2 + z^2 + 2xyz$ at (0, 0, 0)

   (ii) $8z + 2x^2 + 3y^2 + 4z^2 - 3xy$ at (0, 0, –1)

   (iii) $x^4 + y^4 + z^4 - 4xyz$ at (1, 1, 1).

2. Show that the function

$$f(x, y, z) = 2xyz - 4zx - 2yz + x^2 + y^2 + z^2 - 2x - 4y + 4z$$

has 5 stationary points but has a minimum value only at (1, 2, 0).

3. Show that the function

$$3 \log (x^2 + y^2 + z^2) - 2x^3 - 2y^3 - 2z^3,\ (x, y, z) \neq (0, 0, 0)$$

has only one extreme value at $\log (3/e^2)$.

4. Show that the function

$$(y + z)^2 + (z + x)^2 + xyz$$

has no maximum or minimum value.

5. If all the letters are denoted by positive numbers, show that the maximum value of

$$xy(z - h)\left(\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2}\right) \text{ is } (2h/5)^5 ab/c^4.$$

16 Implicit Functions

1. DEFINITION

It is generally assumed that a functional equation $f(x, y) = 0$ determines y as a function of x, but such an equation may not define any such function or it may define one or more than one such function. For example, the equation

$$x^2 + y^2 - 5 = 0$$

determines two functions

$$y = \sqrt{5 - x^2}\,, \text{ and } y = -\sqrt{5 - x^2}\,, \text{ for } x^2 \le 5,$$

whereas the equation

$$x^2 + y^2 + 5 = 0$$

determines no such function.

Definition. Let $f(x, y)$ be a function of two variables, and $y = \phi(x)$ be a function of x such that, for every value of x, for which $\phi(x)$ is defined, $f(x, \phi(x))$ vanishes identically, *i.e.*, $y = \phi(x)$ is a root of the functional equation $f(x, y) = 0$, then $y = \phi(x)$ is an *implicit function* defined by the functional equation $f(x, y) = 0$.

It is only in elementary cases, such as those given above, that it may be possible to express y as a function of x (*i.e.*, determine the implicit function). For more complicated functional equations no such determination of implicit function is possible. The difficulty of actual determination of an analytical expression does not rule out the possibility of the *existence* of the implicit function or functions, defined by a functional equation; the actual determination may demand new processes or may be, from a practical standpoint, too laborious. We now consider an Existence theorem, known as *Implicit function theorem*, that specifies conditions which guarantee that a functional equation does define an implicit function even though actual determination may not be possible. For many purposes, however, it is the fact that an equation does define a function, rather than an expression for the implicit function thus defined, that is of real importance; hence the value of Existence theorem.

1.1 Existence Theorem (*Case of two variables*)

Let $f(x, y)$ be a function of two variables x and y, and (a, b) be a point of its domain of definition such that

(*i*) $f(a, b) = 0$,

(*ii*) *the partial derivatives f_x and f_y exist, and are continuous in a certain neighbourhood of (a, b), and*

(*iii*) $f_y(a, b) \neq 0$, *then there exists a rectangle* $(a-h, a+h; b-k, b+k)$ *about* (a, b) *such that for every value of* x *in the interval* $[a-h, a+h]$, *the equation* $f(x, y) = 0$ *determines one and only one value* $y = \phi(x)$, *lying in the interval* $[b-k, b+k]$, *with the following properties*:

(1) $b = \phi(a)$,

(2) $f[x, \phi(x)] = 0$, *for every* x *in* $[a-h, a+h]$, *and*

(3) $\phi(x)$ *is derivable, and both* $\phi(x)$ *and* $\phi'(x)$ *are continuous in* $[a-h, a+h]$.

Existence. Let f_x, f_y be continuous in a neighbourhood

$$R_1 : (a-h_1, a+h_1; b-k_1, b+k_1), \text{ of } (a, b)$$

Since f_x, f_y exist and are continuous in R_1, therefore f is differentiable and hence continuous in R_1.

Again, since f_y is continuous, and $f_y(a, b) \neq 0$, there exists a rectangle

$$R_2 : (a-h_2, a+h_2; b-k_2, b+k_2),\ h_2 < h_1,\ k_2 < k_1$$

(R_2 contained in R_1) such that for every point of this rectangle, $f_y \neq 0$.

Since $f = 0$ and $f_y \neq 0$ (it is therefore either positive or negative) at the point (a, b), a positive number $k < k_2$ can be found such that

$$f(a, b-k), f(a, b+k)$$

are of opposite signs, for, f is either an increasing or a decreasing function of y, when $y = b$.

Again, since f is continuous, a positive number $h < h_2$ can be found such that for all x in $[a-h, a+h]$,

$$f(x, b-k), f(x, b+k),$$

respectively, may be as near as we please to $f(a, b-k), f(a, b+k)$ and therefore have opposite signs.

Thus, for all x in $[a-h, a+h]$, f is a continuous function of y and changes sign as y changes from $b-k$ to $b+k$. Therefore it vanishes for some y in $[b-k, b+k]$.

Thus, for each x in $[a-h, a+h]$, there is a y in $[b-k, b+k]$ for which $f(x, y) = 0$; this y is a function of x, say $\phi(x)$ such that properties (1) and (2) are true.

Uniqueness. We, now, show that $y = \phi(x)$ is a unique solution of $f(x, y) = 0$ in $R_3 : (a-h, a+h; b-k, b+k)$, that is, $f(x, y)$ cannot be zero for more than one value of y in $[b-k, b+k]$.

Let, if possible, there be two such values y_1, y_2 in $[b-k, b+k]$ so that $f(x, y_1) = 0, f(x, y_2) = 0$. Also $f(x, y)$ considered as a function of a single variable y is derivable in $[b-k, b+k]$, so that by Rolle's theorem, $f_y = 0$ for a value of y between y_1 and y_2, which contradicts the fact that $f_y \neq 0$ in $R_2 \supset R_3$. Hence our supposition is wrong and there cannot be more than one such y.

Derivability. Let (x, y), $(x+\delta x, y+\delta y)$ be two points in $R_3 : (a-h, a+h; b-k, b+k)$ such that

$$y = \phi(x),\ y + \delta y = \phi(x + \delta x)$$

and

$$f(x, y) = 0,\ f(x+\delta x, y+\delta y) = 0$$

Since f is differentiable in R_1 and consequently in R_3 (contained in R_1),

$$\begin{aligned} \therefore \quad 0 &= f(x+\delta x, y+\delta y) - f(x, y) \\ &= \delta x f_x + \delta y f_y + \delta x \psi_1 + \delta y \psi_2 \end{aligned}$$

where ψ_1, ψ_2 are functions of δx and δy, and tend to 0 as

$$(\delta x, \delta y) \to (0, 0)$$

or

$$\frac{\delta y}{\delta x} = -\frac{f_x}{f_y} - \frac{\psi_1}{f_y} - \frac{\delta y}{\delta x}\frac{\psi_2}{f_y} \qquad (f_y \neq 0 \text{ in } R_3)$$

Proceeding to limits as $(\delta x, \delta y) \to (0, 0)$, we get

$$\phi'(x) = \frac{dy}{dx} = -\frac{f_x}{f_y}$$

Thus, $\phi(x)$ is derivable and hence continuous in R_3. Also $\phi'(x)$, being a quotient of two continuous functions, is itself continuous in R_3.

Note: It should be clearly understood that the theorem is essentially of a *local character*. That is, the implicit function $y = \phi(x)$ is a unique solution of $f(x, y) = 0$ in a certain neighbourhood $(a - h, a + h; b - k, b + k)$ of (a, b).

It may have a different solution $y = \psi(x)$ if a different neighbourhood of (a, b) is considered.

Example 1. Let $f(x, y) = x^2 + y^2 - 1$, and a point $(0, 1)$ so that

$$f(0, 1) = 0 \text{ and } f_y(0, 1) \neq 0$$

Of the two possible solutions $y = \pm\sqrt{1 - x^2}$.

(*i*) $y = +\sqrt{1 - x^2}$ is the implicit function in a *nbd* of $(0, 1)$, where $|x| < 1, y > 0$.

(*ii*) $y = -\sqrt{1 - x^2}$ is the implicit function in a *nbd* of $(0, -1)$, where $|x| < 1, y < 0$.

Example 2. Let $f(x, y) = y^2 - yx^2 - 2x^5$, and a point $(1, -1)$ so that

$$f(1, -1) = 0, \text{ and } f_y(1, -1) \neq 0.$$

It can be easily verified that the partial derivatives

$$f_x(x, y) = -2xy - 10x^4, \text{ and } f_y(x, y) = 2y - x^2$$

are continuous in a *nbd* of $(1, -1)$.

Of the two possible solutions

$$y = \frac{x^2}{2}(1 \pm \sqrt{1 + 8x}), \; x > -1/8,$$

$$y = \frac{x^2}{2}(1 - \sqrt{1 + 8x}), \; x > -1/8,$$

is the unique solution of $f(x, y) = 0$ in a *nbd* of $(1, -1)$, since, $-1 = y(1)$.

1.2 General Case

Let $f(x_1, x_2, ..., x_n, y)$ *be a function of* $(n + 1)$ *variables,* $x_1, x_2, ..., x_n, y$ *and* $(a_1, a_2, ..., a_n, b)$ *be a point of its domain of definition such that*

(*i*) $f(a_1, a_2, ..., a_n, b) = 0$

(*ii*) *the partial derivatives with respect to all the* $(n + 1)$ *variables exist and are continuous in a certain neighbourhood of* $(a_1, a_2, ..., a_n, b)$, *and*

(*iii*) $f_y(a_1, a_2, ..., a_n, b) \neq 0$,

then there exists a neighbourhood

$$(a_1 - h_1, a_1 + h_1; a_2 - h_2, a_2 + h_2; ...; a_n - h_n, a_n + h_n; b - k, b + k)$$

of $(a_1, a_2, ..., a_n, b)$ *such that for every point* $(x_1, x_2, ..., x_n)$ *of the neighbourhood*

$$R : (a_1 - h_1, a_1 + h_1; a_2 - h_2, a_2 + h_2; ...; a_n - h_n, a_n + h_n)$$

the equation $f(x_1, x_2, ..., x_n, y) = 0$ *determines one and only one value* $y = \phi(x_1, x_2, ..., x_n)$ *lying in* $[b - k, b + k]$ *with the following properties*:

(1) $b = \phi(a_1, a_2, ..., a_n)$.

(2) $f(x_1, x_2, ..., x_n, \phi) = 0$ *for every point* $(x_1, x_2, ..., x_n)$ *in* R,

(3) ϕ *is continuous and possesses continuous partial derivatives of the first order with respect to* $x_1, x_2, ..., x_n$ *in* R.

The proof is essentially a repetition of that given for the preceding theorem and offers no fresh difficulties.

Ex. Examine the following equations for the existence of unique implicit function near the points indicated and verify by direct calculation. Also find the first derivatives of the solutions whenever these exist.

1. $y^2 + 2x^2y + x^5 = 0, (1, -1)$
2. $y^2 + yx^3 + x^2 = 0, (0, 0)$
3. $y^3 + x^3 - 3xy + y = 0\ (0, 0)$
4. $2xy - \log xy = 2, (1, 1)$
5. $y^4 + y^2x^2 - 2x^5 = 0, (1, 1)$
6. $y^2 - yx^2 - 2x^5 = 0, (0, 0)$
7. $x^2 + xy + y^2 - 1 = 0, (1, 0)$
8. $x_1x_2 - y \log x_2 + \exp(x_1, y) - 1 = 0, (0, 1, 1)$
9. $y^3 \cos x + y^2 \sin^2 x - 7 = 0, (\pi/3, 2)$
10. $xy \sin x + \cos y = 0, (0, \pi/2)$
11. $ax + by + Ax^2 + Hxy + By^2 = 0, (0, 0) (b \neq 0)$

1.3 Derivative of Implicit Functions

When the equation $f(x, y) = 0$ defines y as a function of x that has a derivative dy/dx, that derivative may be obtained simply by differentiating the equation with respect to x, on the understanding that y is a function $\phi(x)$ of x. Thus

$$f_x + f_y \frac{dy}{dx} = 0 \quad ...(1)$$

If the higher partial derivatives of $f(x, y)$ are continuous, we obtain the higher derivatives of y or $\phi(x)$ by successive differentiation of (1), provided always that f_y is not zero; thus

$$f_{xx} + f_{xy}\frac{dy}{dx} + \left(f_{xy} + f_{yy}\frac{dy}{dx}\right)\frac{dy}{dx} + f_y\frac{d^2y}{dx^2} = 0$$

or

$$f_{xx} + 2f_{xy}\frac{dy}{dx} + f_{yy}\left(\frac{dy}{dx}\right)^2 + f_y\frac{d^2y}{dx^2} = 0 \qquad ...(2)$$

Provided f_y is not zero, this equation determines the second derivative, and in a similar way the third and higher derivatives may be found.

Note: If $f(x, y) = 0$, and y is a function of x, we have from (1)

$$\frac{dy}{dx} = \frac{-p}{q}$$

and also from (2),

$$\frac{d^2y}{dx^2} = -\frac{r + 2s(-p/q) + t(-p/q)^2}{q} = \frac{-(rq^2 - 2spq + tp^2)}{q^3}$$

2. JACOBIANS

For further development of the subject, acquaintance with the notion of Jacobians is necessary. We shall now define a Jacobian and also prove some of its important properties.

If $u_1, u_2, ..., u_n$ be n differentiable functions of n variables $x_1, x_2, ..., x_n$, then the determinant

$$\begin{vmatrix} \frac{\partial u_1}{\partial x_1} & \frac{\partial u_1}{\partial x_2} & \cdots & \frac{\partial u_1}{\partial x_n} \\ \frac{\partial u_2}{\partial x_1} & \frac{\partial u_2}{\partial x_2} & \cdots & \frac{\partial u_2}{\partial x_n} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial u_n}{\partial x_1} & \frac{\partial u_n}{\partial x_2} & \cdots & \frac{\partial u_n}{\partial x_n} \end{vmatrix}$$

is called the *Jacobian* or the *Functional Determinant* of the functions $u_1, u_2, ..., u_n$ with respect to $x_1, x_2, ..., x_n$ and is denoted by

$$\frac{\partial(u_1, u_2, ..., u_n)}{\partial(x_1, x_2, ..., x_n)} \text{ or } J\left(\frac{u_1, u_2, ..., u_n}{x_1, x_2, ..., x_n}\right)$$

2.1 Some Properties

Jacobians have the remarkable property of behaving like the derivatives of functions of one variable. A few of the important relations are given here and the proofs depend upon the algebra of determinants.

For $n = 1$, the determinant is simply $\dfrac{\partial y_1}{\partial x_1}$ or $\dfrac{dy_1}{dx_1}$, the derivative of y_1 with respect to x_1; the first of the notations for a Jacobian is suggested by a certain analogy between the properties of the Jacobian and the derivative.

Theorem 1. *If $u_1, u_2, ..., u_n$ are functions of $y_1, y_2, ..., y_n$ and $y_1, y_2, ..., y_n$ are themselves functions of $x_1, x_2, ..., x_n$, then*

$$\frac{\partial(u_1, u_2, ..., u_n)}{\partial(x_1, x_2, ..., x_n)} = \frac{\partial(u_1, u_2, ..., u_n)}{\partial(y_1, y_2, ..., y_n)} \frac{\partial(y_1, y_2, ..., y_n)}{\partial(x_1, x_2, ..., x_n)} \quad ...(1)$$

For $n = 1$, the theorem reduces to the usual notation

$$\frac{du_1}{dx_1} = \frac{du_1}{dy_1} \frac{dy_1}{dx_1}$$

The proof of the theorem depends on the "row by column" rule of multiplication of determinants combined with the rule for the derivative of a function of a function.

Thus for determinants on the right hand side of (1), rth row of the first is $\dfrac{\partial u_r}{\partial y_1}, \dfrac{\partial u_r}{\partial y_2}, ..., \dfrac{\partial u_r}{\partial y_n}$, sth column of the second is $\dfrac{\partial y_1}{\partial x_s}, \dfrac{\partial y_2}{\partial x_s}, ..., \dfrac{\partial y_n}{\partial x_s}$, so that the element in the rth row and the sth column of the product is

$$\frac{\partial u_r}{\partial y_1}\frac{\partial y_1}{\partial x_s} + \frac{\partial u_r}{\partial y_2}\frac{\partial y_2}{\partial x_s} + ... + \frac{\partial u_r}{\partial y_n}\frac{\partial y_n}{\partial x_s}$$

and this is equal to $\dfrac{\partial u_r}{\partial x_s}$, which is the element in the rth row and the sth column of the Jacobian on the left hand side. Hence the theorem.

Corollary. If $x_r = u_r$, $r = 1, 2, ..., n$ and assuming the existence of inverse functions $x_1, x_2, ..., x_n$ (that is, assuming that the equations which define $y_1, y_2, ..., y_n$ as functions of $x_1, x_2, ..., x_n$ determine $x_1, x_2, ..., x_n$ as functions of $y_1, y_2, ..., y_n$) we find

$$\frac{\partial(x_1, x_2, ..., x_n)}{\partial(y_1, y_2, ..., y_n)} \cdot \frac{\partial(y_1, y_2, ..., y_n)}{\partial(x_1, x_2, ..., x_n)} = \frac{\partial(x_1, x_2, ..., x_n)}{\partial(x_1, x_2, ..., x_n)} = 1 \quad ...(2)$$

since $\dfrac{\partial x_i}{\partial x_j} = 0$, for $i \neq j = 1$, for $i = j$

Theorem 2. *If $y_1, y_2, ..., y_n$ are determined as functions of $x_1, x_2, ..., x_n$ by the equations*

$$\phi_r(x_1, x_2, ..., x_n, y_1, y_2, ..., y_n) = 0,\ r = 1, 2, ..., n$$

then

$$\frac{\partial(\phi_1, \phi_2, \ldots, \phi_n)}{\partial(x_1, x_2, \ldots, x_n)} = (-1)^n \frac{\partial(\phi_1, \phi_2, \ldots, \phi_n)}{\partial(y_1, y_2, \ldots, y_n)} \frac{\partial(y_1, y_2, \ldots, y_n)}{\partial(x_1, x_2, \ldots, x_n)} \qquad \ldots(3)$$

[Theorem 1 is a particular form of this theorem.]

Differentiating the equations $\phi_r = 0$ with respect to x_s, we get

$$\frac{\partial \phi_r}{\partial x_s} + \frac{\partial \phi_r}{\partial y_1}\frac{\partial y_1}{\partial x_s} + \frac{\partial \phi_r}{\partial y_2}\frac{\partial y_2}{\partial x_s} + \ldots + \frac{\partial \phi_r}{\partial y_n}\frac{\partial y_n}{\partial x_s} = 0$$

or

$$\frac{\partial \phi_r}{\partial y_1}\cdot\frac{\partial y_1}{\partial x_s} + \frac{\partial \phi_r}{\partial y_2}\frac{\partial y_2}{\partial x_s} + \ldots + \frac{\partial \phi_r}{\partial y_n}\frac{\partial y_n}{\partial x_s} = -\frac{\partial \phi_r}{\partial x_s}$$

so that the element in the rth row and the sth column of the determinant which is the product of the two determinants on the right of (3) is $-\dfrac{\partial \phi_r}{\partial x_s}$, from which the result follows.

Theorem 3. *(i) If $y_{m+1}, y_{m+2}, \ldots, y_n$ are constant with respect to $x_1, x_2, \ldots, x_m$, or (ii) if $y_1, y_2, \ldots, y_m$ are constant with respect to $x_{m+1}, x_{m+2}, \ldots, x_n$, then*

$$\frac{\partial(y_1, y_2, \ldots, y_m, \ldots, y_n)}{\partial(x_1, x_2, \ldots, x_m, \ldots, x_n)} = \frac{\partial(y_1, y_2, \ldots, y_m)}{\partial(x_1, x_2, \ldots, x_m)} \cdot \frac{\partial(y_{m+1}, y_{m+2}, \ldots, y_n)}{\partial(x_{m+1}, x_{m+2}, \ldots, x_n)} \qquad \ldots(4)$$

(*i*) $\dfrac{\partial y_r}{\partial x_s} = 0$, when $r = m+1, m+2, \ldots, n; s = 1, 2, \ldots, m$.

$$\frac{\partial(y_1, y_2, \ldots, y_n)}{\partial(x_1, x_2, \ldots, x_n)} = \begin{vmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_1}{\partial x_2} & \cdots \frac{\partial y_1}{\partial x_m} & \frac{\partial y_1}{\partial x_{m+1}} & \cdots \frac{\partial y_1}{\partial x_n} \\ \frac{\partial y_2}{\partial x_1} & \frac{\partial y_2}{\partial x_2} & \cdots \frac{\partial y_2}{\partial x_m} & \frac{\partial y_2}{\partial x_{m+1}} & \cdots \frac{\partial y_2}{\partial x_n} \\ \vdots & \vdots & \vdots \;\vdots & \vdots & \vdots\;\vdots \\ \frac{\partial y_m}{\partial x_1} & \frac{\partial y_m}{\partial x_2} & \cdots \frac{\partial y_m}{\partial x_m} & \frac{\partial y_m}{\partial x_{m+1}} & \cdots \frac{\partial y_m}{\partial x_n} \\ \frac{\partial y_{m+1}}{\partial x_1} & \frac{\partial y_{m+1}}{\partial x_2} & \cdots \frac{\partial y_{m+1}}{\partial x_m} & \frac{\partial y_{m+1}}{\partial x_{m+1}} & \cdots \frac{\partial y_{m+1}}{\partial x_n} \\ \vdots & \vdots & \vdots \;\vdots & \vdots & \vdots\;\vdots \\ \frac{\partial y_n}{\partial x_1} & \frac{\partial y_n}{\partial x_2} & \cdots \frac{\partial y_n}{\partial x_m} & \frac{\partial y_n}{\partial x_{m+1}} & \cdots \frac{\partial y_n}{\partial x_n} \end{vmatrix}$$

$$= \begin{vmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_1}{\partial x_2} & \cdots & \frac{\partial y_1}{\partial x_m} & \frac{\partial y_1}{\partial x_{m+1}} & \cdots & \frac{\partial y_m}{\partial x_n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{\partial y_m}{\partial x_1} & \frac{\partial y_m}{\partial x_2} & \cdots & \frac{\partial y_m}{\partial x_m} & \frac{\partial y_m}{\partial x_{m+1}} & \cdots & \frac{\partial y_m}{\partial x_n} \\ 0 & 0 & \cdots & 0 & \frac{\partial y_{m+1}}{\partial x_{m+1}} & \cdots & \frac{\partial y_{m+1}}{\partial x_n} \\ 0 & 0 & \cdots & 0 & \frac{\partial y_{m+2}}{\partial x_{m+1}} & \cdots & \frac{\partial y_{m+2}}{\partial x_n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & \frac{\partial y_n}{\partial x_{m+1}} & \cdots & \frac{\partial y_n}{\partial x_n} \end{vmatrix}$$

$$= \frac{\partial(y_1, y_2, ..., y_m)}{\partial(x_1, x_2, ..., x_n)} \cdot \frac{\partial(y_{m+1}, y_{m+2}, ..., y_n)}{\partial(x_{m+1}, x_{m+2}, ..., x_n)}$$

(*ii*) may also be proved similarly.

Corollary. *In particular,*

$$\frac{\partial(y_1, ..., y_m, x_{m+1}, ..., x_n)}{\partial(x_1, ..., x_m, x_{m+1}, ..., x_n)} = \frac{\partial(y_1, ..., y_m)}{\partial(x_1, ..., x_m)} \qquad ...(5)$$

Theorem 4. *If u, v are functions of ξ, η, ζ, and the variables ξ, η, ζ, are themselves functions of the independent variables x and y, then*

$$\frac{\partial(u, v)}{\partial(x, y)} = \frac{\partial(u, v)}{\partial(\xi, \eta)} \cdot \frac{\partial(\xi, \eta)}{\partial(x, y)} + \frac{\partial(u, v)}{\partial(\eta, \xi)} \cdot \frac{\partial(\eta, \xi)}{\partial(x, y)} + \frac{\partial(u, v)}{\partial(\zeta, \xi)} \cdot \frac{\partial(\zeta, \xi)}{\partial(x, y)} \qquad ...(6)$$

We have

$$\frac{\partial u}{\partial x} = \frac{\partial u}{\partial \xi}\frac{\partial \xi}{\partial x} + \frac{\partial u}{\partial \eta}\frac{\partial \eta}{\partial x} + \frac{\partial u}{\partial \zeta}\frac{\partial \zeta}{\partial x} \qquad ...(7)$$

$$\frac{\partial u}{\partial y} = \frac{\partial u}{\partial \xi}\frac{\partial \xi}{\partial y} + \frac{\partial u}{\partial \eta}\frac{\partial \eta}{\partial y} + \frac{\partial u}{\partial \zeta}\frac{\partial \zeta}{\partial y} \qquad ...(8)$$

and if we substitute these values in the Jacobian $\frac{\partial(u, v)}{\partial(x, y)}$, we get

$$\frac{\partial(u, v)}{\partial(x, y)} = \frac{\partial u}{\partial \xi}\frac{\partial(\xi, v)}{\partial(x, y)} + \frac{\partial u}{\partial \eta}\frac{\partial(\eta, v)}{\partial(x, y)} + \frac{\partial u}{\partial \zeta}\frac{\partial(\zeta, v)}{\partial(x, y)} \qquad ...(9)$$

which is a linear expression of the Jacobians of (ξ, v), (η, v) and (ζ, v) with respect to x and y.

Now in each Jacobian on the right of equation (9), substitute the expressions for $\frac{\partial v}{\partial x}$ and $\frac{\partial v}{\partial y}$ which are similar to (7) and (8). Each of these Jacobians will be given as a linear expression of the Jacobians of (ξ, η), (η, ζ) and (ζ, ξ) since those of (ξ, ξ), (η, η) and (ζ, ζ) have two identical parallel lines and so vanish. Thus we see that the terms which involve the Jacobian of (ξ, η) are

$$\frac{\partial u}{\partial \xi}\frac{\partial v}{\partial \eta}\frac{\partial(\xi, \eta)}{\partial(x, y)} + \frac{\partial u}{\partial \eta}\frac{\partial v}{\partial \xi}\frac{\partial(\eta, \xi)}{\partial(x, y)}$$

which is equal to $\frac{\partial(u, v)}{\partial(\xi, \eta)}\frac{\partial(\xi, \eta)}{\partial(x, y)}$, the first terms on the right of equation (6).

Similarly we obtain the remaining two terms and the formula is established.

Example 3. If $x = r\sin\theta\cos\phi$, $y = r\sin\theta\sin\phi$, $z = r\cos\theta$, then show that

$$\frac{\partial(x, y, z)}{\partial(r, \theta, \phi)} = r^2\sin\theta$$

■
$$\frac{\partial(x, y, z)}{\partial(r, \theta, \phi)} = \begin{vmatrix} \sin\theta\cos\phi & r\cos\theta\cos\phi & -r\sin\theta\sin\phi \\ \sin\theta\sin\phi & r\cos\theta\sin\phi & r\sin\theta\cos\phi \\ \cos\theta & -r\sin\theta & 0 \end{vmatrix}$$

$$= r^2\sin\theta \begin{vmatrix} \sin\theta\cos\phi & \cos\theta\cos\phi & -\sin\theta \\ \sin\theta\sin\phi & \cos\theta\sin\phi & \cos\phi \\ \cos\theta & -\sin\theta & 0 \end{vmatrix}$$

Adding $(\cos\phi)\ R_1$ to $(\sin\phi)\ R_2$,

$$= \frac{r^2\sin\theta}{\sin\phi} \begin{vmatrix} \sin\theta\cos\phi & \cos\theta\cos\phi & -\sin\phi \\ \sin\theta & \cos\theta & 0 \\ \cos\theta & -\sin\theta & 0 \end{vmatrix} = r^2\sin\theta$$

Example 4. If $y_1 + y_2 + \ldots + y_n = x_1$, $y_2 + y_3 + \ldots + y_n = x_1x_2$, ..., $y_r + y_{r+1} + \ldots + y_n = x_1x_2 \ldots x_r$, ..., $y_n = x_1x_2 \ldots x_n$, then show that

$$\frac{\partial(y_1, y_2, \ldots, y_n)}{\partial(x_1, x_2, \ldots, x_n)} = x_1^{n-1}x_2^{n-2}\ldots x_{n-2}^2 x_{n-1}.$$

■ Solving for $y_1, y_2, \ldots, y_n$, we get

$$\begin{aligned} y_1 &= x_1 - x_1x_2 = x_1\,(1 - x_2) \\ y_2 &= x_1x_2 - x_1x_2x_3 = x_1x_2\,(1 - x_3) \\ &\vdots \\ y_{n-1} &= x_1x_2 \ldots x_{n-1}\,(1 - x_n) \\ y_n &= x_1x_2 \ldots x_n \end{aligned}$$

$$\therefore \quad \frac{\partial(y_1, y_2, ..., y_n)}{\partial(x_1, x_2, ..., x_n)} =$$

$$\begin{vmatrix} 1-x_2 & -x_1 & 0 & \dots & 0 \\ x_2(1-x_3) & x_1(1-x_3) & -x_1x_2 & \dots & 0 \\ x_2x_3(1-x_4) & x_1x_3(1-x_4) & x_1x_2(1-x_4) & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ x_2x_3...x_{n-1}(1-x_n) & x_1x_3...x_{n-1}(1-x_n) & x_1x_2x_4...x_{n-1}(1-x_n) & \dots & x_1x_2...x_{n-1} \\ x_3x_3...x_n & x_1x_2x_4...x_n & x_1x_3...x_n & \dots & x_1x_2...x_{n-1} \end{vmatrix}$$

Adding R_n to R_{n-1}, then R_{n-1} to R_{n-2}, ..., then R_2 to R_1 and expanding by last column

$$= (x_1x_2...x_{n-1})\,(x_1x_2...x_{n-2})...(x_1x_2)(x_1)$$

$$= x_1^{n-1}x_2^{n-2}...x_{n-2}^2x_{n-1}$$

Example 5. The roots of the equation in λ

$$(\lambda - x)^3 + (\lambda - y)^3 + (\lambda - z)^3 = 0$$

are u, v, w. Prove that

$$\frac{\partial(u, v, w)}{\partial(x, y, z)} = -2\frac{(y-z)(z-x)(x-y)}{(v-w)(w-u)(u-v)}$$

■ Here u, v, w are roots of the equation

$$\lambda^3 - (x+y+z)\lambda^2 + (x^2+y^2+z^2)\lambda - \frac{1}{3}(x^3+y^3+z^3) = 0$$

Let
$$x+y+z = \xi,\ x^2+y^2+z^2 = \eta,\ \frac{1}{3}(x^3+y^3+z^3) = \zeta \qquad ...(1)$$

and

$$u+v+w = \xi,\ vw+wu+uv = \eta,\ uvw = \zeta \qquad ...(2)$$

Hence from (1),

$$\frac{\partial(\xi, \eta, \zeta)}{\partial(x, y, z)} = \begin{vmatrix} 1 & 1 & 1 \\ 2x & 2y & 2z \\ x^2 & y^2 & z^2 \end{vmatrix}$$

$$= 2(y-z)\,(z-x)\,(x-y) \qquad ...(3)$$

and from (2),

$$\frac{\partial(\xi, \eta, \zeta)}{\partial(u, v, w)} = \begin{vmatrix} 1 & 1 & 1 \\ v+w & w+u & u+v \\ vw & wu & uv \end{vmatrix}$$

$$= -(v-w)\,(w-u)\,(u-v) \qquad ...(4)$$

Hence from (3) and (4) and using theorem 1, we get

$$\frac{\partial(u, v, w)}{\partial(x, y, z)} = \frac{\partial(u, v, w)}{\partial(\xi, \eta, \zeta)} \cdot \frac{\partial(\xi, \eta, \zeta)}{\partial(x, y, z)} = -2\frac{(y-z)(z-x)(x-y)}{(v-w)(w-u)(u-v)}$$

Example 6. If $y_r = \dfrac{u_r}{u}$, $r = 1, 2, \ldots, n$, and if u and u_r are functions of the n independent variables x_1, $x_2, \ldots, x_n$, prove that

$$\frac{\partial(y_1, y_2, \ldots, y_n)}{\partial(x_1, x_2, \ldots, x_n)} = \frac{1}{u^{n+1}} \begin{vmatrix} u & \dfrac{\partial u}{\partial x_1} & \dfrac{\partial u}{\partial x_2} & \cdots & \dfrac{\partial u}{\partial x_n} \\ u_1 & \dfrac{\partial u_1}{\partial x_1} & \dfrac{\partial u_1}{\partial x_2} & \cdots & \dfrac{\partial u_1}{\partial x_n} \\ u_2 & \dfrac{\partial u_2}{\partial x_1} & \dfrac{\partial u_2}{\partial x_2} & \cdots & \dfrac{\partial u_2}{\partial x_n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ u_n & \dfrac{\partial u_n}{\partial x_1} & \dfrac{\partial u_n}{\partial x_2} & \cdots & \dfrac{\partial u_n}{\partial x_n} \end{vmatrix}$$

■ Now

$$\frac{\partial y_r}{\partial x_s} = \frac{1}{u}\frac{\partial u_r}{\partial x_s} - \frac{u_r}{u^2}\frac{\partial u}{\partial x_s}$$

$$\therefore \quad \frac{\partial(y_1, y_2, \ldots, y_n)}{\partial(x_1, x_2, \ldots, x_n)} = \begin{vmatrix} \dfrac{1}{u}\dfrac{\partial u_1}{\partial x_1} - \dfrac{u_1}{u^2}\dfrac{\partial u}{\partial x_1} & \cdots & \dfrac{1}{u}\dfrac{\partial u_1}{\partial x_n} - \dfrac{u_1}{u^2}\dfrac{\partial u}{\partial x_n} \\ \dfrac{1}{u}\dfrac{\partial u_2}{\partial x_1} - \dfrac{u_2}{u^2}\dfrac{\partial u}{\partial x_1} & \cdots & \dfrac{1}{u}\dfrac{\partial u_2}{\partial x_n} - \dfrac{u_2}{u^2}\dfrac{\partial u}{\partial x_n} \\ \vdots & \vdots & \vdots \\ \dfrac{1}{u}\dfrac{\partial u_n}{\partial x_1} - \dfrac{u_n}{u^2}\dfrac{\partial u}{\partial x_1} & \cdots & \dfrac{1}{u}\dfrac{\partial u_n}{\partial x_n} - \dfrac{u_n}{u^2}\dfrac{\partial u}{\partial x_n} \end{vmatrix}$$

Taking out $\dfrac{1}{u}$ from each column and bordering the determinant, we get

$$\frac{\partial(y_1, y_2, \ldots, y_n)}{\partial(x_1, x_2, \ldots, x_n)} = \frac{1}{u^n} \begin{vmatrix} 1 & 0 & \cdots & 0 \\ u_1 & \dfrac{\partial u_1}{\partial x_1} - \dfrac{u_1}{u}\dfrac{\partial u}{\partial x_1} & \cdots & \dfrac{\partial u_1}{\partial x_n} - \dfrac{u_1}{u}\dfrac{\partial u}{\partial x_n} \\ u_2 & \dfrac{\partial u_2}{\partial x_1} - \dfrac{u_2}{u}\dfrac{\partial u}{\partial x_1} & \cdots & \dfrac{\partial u_2}{\partial x_n} - \dfrac{u_2}{u}\dfrac{\partial u}{\partial x_n} \\ \vdots & \vdots & \vdots & \vdots \\ u_n & \dfrac{\partial u_n}{\partial x_1} - \dfrac{u_n}{u}\dfrac{\partial u}{\partial x_1} & \cdots & \dfrac{\partial u_n}{\partial x_n} - \dfrac{u_n}{u}\dfrac{\partial u}{\partial x_n} \end{vmatrix}$$

Operating $\quad C_2 + \frac{1}{u}\frac{\partial u}{\partial x_1}C_1,\ C_3 + \frac{1}{u}\frac{\partial u}{\partial x_2}C_1,\ \ldots,\ C_{n+1} + \frac{1}{u}\frac{\partial u}{\partial x_n}C_1$

$$= \frac{1}{u^n}\begin{vmatrix} 1 & \frac{1}{u}\frac{\partial u}{\partial x_1} & \cdots & \frac{1}{u}\frac{\partial u}{\partial x_n} \\ u_1 & \frac{\partial u_1}{\partial x_1} & \cdots & \frac{\partial u_1}{\partial x_n} \\ u_2 & \frac{\partial u_2}{\partial x_1} & \cdots & \frac{\partial u_2}{\partial x_n} \\ \vdots & \vdots & \vdots & \vdots \\ u_n & \frac{\partial u_n}{\partial x_1} & \cdots & \frac{\partial u_n}{\partial x_n} \end{vmatrix}$$

$$= \frac{1}{u^{n-1}}\begin{vmatrix} u & \frac{\partial u}{\partial x_1} & \frac{\partial u}{\partial x_2} & \cdots & \frac{\partial u}{\partial x_n} \\ u_1 & \frac{\partial u_1}{\partial x_1} & \frac{\partial u_1}{\partial x_2} & \cdots & \frac{\partial u_1}{\partial x_n} \\ u_2 & \frac{\partial u_2}{\partial x_1} & \frac{\partial u_2}{\partial x_2} & \cdots & \frac{\partial u_2}{\partial x_n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ u_n & \frac{\partial u_n}{\partial x_1} & \frac{\partial u_n}{\partial x_2} & \cdots & \frac{\partial u_n}{\partial x_n} \end{vmatrix}$$

Example 7. If $u = \frac{x^2 + y^2 + z^2}{x}$, $v = \frac{x^2 + y^2 + z^2}{y}$, and $w = \frac{x^2 + y^2 + z^2}{z}$ find $\frac{\partial(x, y, z)}{\partial(u, v, w)}$.

$$\frac{\partial(u, v, w)}{\partial(x, y, z)} = \begin{vmatrix} 1 - \frac{y^2 + z^2}{x^2} & \frac{2y}{x} & \frac{2z}{x} \\ \frac{2x}{y} & 1 - \frac{x^2 + z^2}{y^2} & \frac{2z}{y} \\ \frac{2x}{z} & \frac{2y}{z} & 1 - \frac{x^2 + y^2}{z^2} \end{vmatrix}$$

- Applying $C_1 \to C_1 + \frac{y}{x}C_2 + \frac{z}{x}C_3$

$$\frac{\partial(u, v, w)}{\partial(x, y, z)} = \begin{vmatrix} \frac{x^2 + y^2 + z^2}{x^2} & \frac{2y}{x} & \frac{2z}{x} \\ \frac{x^2 + y^2 + z^2}{xy} & 1 - \frac{x^2 + z^2}{y^2} & \frac{2z}{y} \\ \frac{x^2 + y^2 + z^2}{xz} & \frac{2z}{z} & 1 - \frac{x^2 + y^2}{z^2} \end{vmatrix}$$

$$= \frac{(x^2+y^2+z^2)}{x^2 \cdot xy \cdot xz} \begin{vmatrix} 1 & 2xy & 2xz \\ 1 & xy - \frac{x}{y}(x^2+z^2) & 2xz \\ 1 & 2yz & xz - \frac{x}{z}(x^2+y^2) \end{vmatrix}$$

$$= \frac{(x^2+y^2+z^2)}{x^4 yz} \begin{vmatrix} 1 & 2xy & 2xz \\ 0 & -\frac{x(x^2+y^2+z^2)}{y} & 0 \\ 0 & 0 & -\frac{x}{z}(x^2+y^2+z^2) \end{vmatrix}$$

$$= \frac{(x^2+y^2+z^2)^3}{x^2y^2z^2}$$

$$\therefore \quad \frac{\partial(x, y, z)}{\partial(u, v, w)} = \frac{x^2y^2z^2}{(x^2+y^2+z^2)^3}$$

Ex. 1. If $u = \cos x$, $v = \sin x \cos y$, $w = \sin x \sin y \cos z$, then show that

$$\frac{\partial(u, v, w)}{\partial(x, y, z)} = (-1)^2 \sin^3 x \sin^2 y \sin z.$$

Ex. 2. If $u = a \cosh x \cos y$, $v = a \sinh x \sin y$, then show that

$$\frac{\partial(u, v)}{\partial(x, y)} = \frac{1}{2}a^2(\cosh 2x - \cos 2y).$$

Ex. 3. If $x + y + z = u$, $y + z = uv$, $z = uvw$, then show that

$$\frac{\partial(x, y, z)}{\partial(u, v, w)} = u^2 v \cdot$$

Ex. 4. If α, β, γ are the roots of the equation in t, such that

$$\frac{u}{a+t} + \frac{v}{b+t} + \frac{w}{c+t} = 1,$$

then prove that

$$\frac{\partial(u, v, w)}{\partial(\alpha, \beta, \gamma)} = -\frac{(\beta-\gamma)(\gamma-\alpha)(\alpha-\beta)}{(b-c)(c-a)(a-b)}.$$

Ex. 5. If $u = x/(1 - r^2)^{1/2}$, $v = y/(1 - r^2)^{1/2}$, $w = z/(1 - r^2)^{1/2}$ where $r^2 = x^2 + y^2 + z^2$, then show that

$$J\left(\begin{matrix} u, v, w \\ x, y, z \end{matrix}\right) = \frac{1}{(1 - r^2)^{5/2}}.$$

3. STATIONARY VALUES UNDER SUBSIDIARY CONDITIONS

To find the stationary values of the function

$$f(x_1, x_2, ..., x_n, u_1, u_2, ..., u_m) \quad ...(1)$$

of $(n + m)$ *variables which are connected by* m *differentiable equations*

$$\phi_r(x_1, x_2, ..., x_n, u_1, u_2, ..., u_m) = 0, \; r = 1, 2, ..., m \quad ...(2)$$

If the m variables $u_1, u_2, ..., u_m$ are determinate as functions of $x_1, x_2, ..., x_n$ from the system of m equations of (2), then f can be regarded as a function of n independent variables $x_1, x_2, ..., x_n$.

At a stationary point of f, $df = 0$.

Hence at a stationary point (by § 11.1, Ch. 15).

$$0 = df = f_{x_1} dx_1 + f_{x_2} dx_2 + ... + f_{x_n} dx_n + f_{u_1} du_1 + ... + f_{u_m} du_m \quad ...(3)$$

Again differentiating the equation (2), we get

$$\left.\begin{aligned} &\frac{\partial \phi_1}{\partial x_1} dx_1 + ... + \frac{\partial \phi_1}{\partial x_n} dx_n + \frac{\partial \phi_1}{\partial u_1} du_1 + ... + \frac{\partial \phi_1}{\partial u_m} du_m = 0 \\ &\frac{\partial \phi_2}{\partial x_1} dx_1 + ... + \frac{\partial \phi_2}{\partial x_n} dx_n + \frac{\partial \phi_2}{\partial u_1} du_1 + ... + \frac{\partial \phi_2}{\partial u_m} du_m = 0 \\ &\vdots \\ &\frac{\partial \phi_m}{\partial x_1} dx_1 + ... + \frac{\partial \phi_m}{\partial x_n} dx_n + \frac{\partial \phi_m}{\partial u_1} du_1 + ... + \frac{\partial \phi_m}{\partial u_m} du_m = 0 \end{aligned}\right\} \quad ...(4)$$

From these m equations of (4), the differentials $du_1, du_2, ..., du_m$ of the m dependent variables may be found in terms of the n differentials $dx_1, dx_2, ..., dx_n$ and substituted in (3). This way df has been expressed in terms of the differentials of the independent variables, and since the differentials of the independent variables are arbitrary and $df = 0$, the coefficients of each of these n differentials may be equated to zero. These n equations together with the m equations of (2) constitute a system of $(n + m)$ equations to determine the $(n + m)$ coordinates of the stationary points of f.

Example 8. $F(x, y, z)$ is a function subject to the constraint condition $G(x, y, z) = 0$. Show that at a stationary point

$$F_x G_y - F_y G_x = 0$$

■ We may consider z as a function of the independent variables x, y.

As a stationary point, $dF = 0$

$$\therefore \quad 0 = dF = F_x dx + F_y dy + F_z dz \quad ...(1)$$

Differentiating the relation $G(x, y, z) = 0$, we get

$$G_x\,dx + G_y\,dy + G_z\,dz = 0 \qquad ...(2)$$

Putting the value of dz from (2) into (1), or what is same thing, eliminating dz from (1) and (2), we get

$$(F_xG_z - G_xF_z)\,dx + (F_yG_z - G_yF_z)\,dy = 0 \qquad ...(3)$$

Since dx, dy (being differentials of independent variables) are arbitrary, therefore

$$F_xG_z - G_xF_z = 0$$
$$F_yG_z - G_yF_z = 0$$

which give

$$F_xG_y - F_yG_x = 0.$$

Ex. 1. Find the stationary points of the function xy^2z^2 subject to the conditions

$$x + y + z = 6,\ x > 0,\ y > 0,\ z > 0.$$

Ex. 2. Find the stationary points of $x^2 + y^2$ subject to the condition

$$3x^2 + 4xy + 6y^2 = 140.$$

Ex. 3. Find the stationary points of the function $x^2y^2z^2$ subject to the condition

$$x^2 + y^2 + z^2 = a^2.$$

Ex. 4. Find the stationary points of the function xyz, when x, y, z are connected by the equation

$$x^2/9 + y^2/16 + z^2/36 = 1.$$

3.1 Lagrange's Undetermined Multipliers

In this section we shall discuss the determination of stationary points from a modified point of view. The process consists in the introduction of undetermined multipliers, a method due to Lagrange.

Before we discuss the method proper, let us notice that in the above section (§ 3) the differentials $du_1, du_2, ..., du_m$ of the m dependent variables were found from equations of (4) and substituted in (3), so as to express df in terms of the differentials of the independent variables. All this amounts to elimination of the differentials of the dependent variables from (3) and (4). This process of elimination is effected due to Lagrange by the introduction of multipliers. The process is as follows:

Multiply each of the equations of (4) by $\lambda_1, \lambda_2, ..., \lambda_m$, which are to be specified later, and add to (3). Now the m multipliers are so chosen that the coefficients of the m differentials of dependent variables all vanish. This gives m equations to determine λ's. Thus in equation (3) there remain only the differentials of n independent variables, the coefficient of each one of them may be equated to zero. The process thus gives $n + m$ equations which along with m constraint conditions form $n + 2m$ equations, to determine the m multipliers $\lambda_1, \lambda_2, ..., \lambda_m$, and $m + n$ variables $x_1, x_2, ..., x_n, u_1, u_2, ..., u_m$, for which the function has a stationary value.

3.2 Lagrange's Method of Multipliers

To find the stationary points of the function

$$f(x_1, x_2, ..., x_n, u_1, u_2, ..., u_m) \qquad ...(1)$$

of $n + m$ variables which are connected by the equations

$$\phi_r(x_1, x_2, ..., x_n, u_1, u_2, ..., u_m) = 0,\ r = 1, 2, ..., m \qquad ...(2)$$

As in § 3, if the m variables $u_1, u_2, ..., u_m$ from equations of (2) are expressed in terms of variables $x_1, x_2, ..., x_n$ the function f may be considered as a function of n independent variables

$$x_1, x_2, ..., x_n$$

For stationary values, $df = 0$

$$\therefore \qquad 0 = df = \frac{\partial f}{\partial x_1}dx_1 + \frac{\partial f}{\partial x_2}dx_2 + ... + \frac{\partial f}{\partial x_n}dx_n + \frac{\partial f}{\partial u_1}du_1 + ... + \frac{\partial f}{\partial u_m}du_m \qquad ...(3)$$

Differentiating equations of (2), we get

$$\left.\begin{aligned} &\frac{\partial \phi_1}{\partial x_1}dx_1 + ... + \frac{\partial \phi_1}{\partial x_n}dx_n + \frac{\partial \phi_1}{\partial u_1}du_1 + ... + \frac{\partial \phi_1}{\partial u_m}du_m = 0 \\ &\frac{\partial \phi_2}{\partial x_1}dx_1 + ... + \frac{\partial \phi_2}{\partial x_n}dx_n + \frac{\partial \phi_2}{\partial u_1}du_1 + ... + \frac{\partial \phi_2}{\partial u_m}du_m = 0 \\ &\vdots \\ &\frac{\partial \phi_m}{\partial x_1}dx_1 + ... + \frac{\partial \phi_m}{\partial x_n}dx_n + \frac{\partial \phi_m}{\partial u_1}du_1 + ... + \frac{\partial \phi_m}{\partial u_m}du_m = 0 \end{aligned}\right\} \qquad ...(4)$$

Multiplying the equations of (4) by $\lambda_1, \lambda_2, ..., \lambda_m$ respectively and adding to the equation (3), we get

$$0 = df = \left(\frac{\partial f}{\partial x_1} + \Sigma\lambda_r \frac{\partial \phi_r}{\partial x_1}\right)dx_1 + ... + \left(\frac{\partial f}{\partial x_n} + \Sigma\lambda_r \frac{\partial \phi_r}{\partial x_n}\right)dx_n$$

$$+ \left(\frac{\partial f}{\partial u_1} + \Sigma\lambda_r \frac{\partial \phi_r}{\partial u_1}\right)du_1 + ... + \left(\frac{\partial f}{\partial u_m} + \Sigma\lambda_r \frac{\partial \phi_r}{\partial u_m}\right)du_m \qquad ...(5)$$

Let the m multipliers $\lambda_1, \lambda_2, ..., \lambda_m$ be so chosen that the coefficients of the m differentials $du_1, du_2, ..., du_m$ all vanish, *i.e.*,

$$\frac{\partial f}{\partial u_1} + \Sigma\lambda_r \frac{\partial \phi_r}{\partial u_1} = 0, ..., \frac{\partial f}{\partial u_m} + \Sigma\lambda_r \frac{\partial \phi_r}{\partial u_m} = 0 \qquad ...(6)$$

Then (5) becomes

$$0 = df = \left(\frac{\partial f}{\partial x_1} + \Sigma\lambda_r \frac{\partial \phi_r}{\partial x_1}\right)dx_1 + ... + \left(\frac{\partial f}{\partial x_n} + \Sigma\lambda_r \frac{\partial \phi_r}{\partial x_n}\right)dx_n$$

so that the differential df is expressed in terms of the differentials of independent variables only. Hence

$$\frac{\partial f}{\partial x_1} + \Sigma\lambda_r \frac{\partial \phi_r}{\partial x_1} = 0, ..., \frac{\partial f}{\partial x_n} + \Sigma\lambda_r \frac{\partial \phi_r}{\partial x_n} = 0 \qquad ...(7)$$

Equations (2), (6) and (7) form a system of $n + 2m$ equations which may be simultaneously solved to determine the m multipliers $\lambda_1, \lambda_2, ..., \lambda_m$ and the $n + m$ coordinates $x_1, x_2, ..., x_n, u_1, u_2, ..., u_m$ of the stationary points of f.

An Important Rule. For practical purposes, the process of obtaining equations (6) and (7) of the above section, may be put in a precise form as follows:

Define a function

$$F = f + \lambda_1\phi_1 + \lambda_2\phi_2 + \ldots + \lambda_m\phi_m$$

and consider all the variables $x_1, x_2, \ldots, x_n, u_1, u_2, \ldots, u_m$ as independent.

At a stationary point of F, $dF = 0$. Therefore,

$$0 = dF = \frac{\partial F}{\partial x_1}dx_1 + \frac{\partial F}{\partial x_2}dx_2 + \ldots + \frac{\partial F}{\partial x_n}dx_n + \frac{\partial F}{\partial u_1}du_1 + \ldots + \frac{\partial F}{\partial u_m}du_m$$

$$\therefore \qquad \frac{\partial F}{\partial x_1} = 0, \ldots, \frac{\partial F}{\partial x_n} = 0, \frac{\partial F}{\partial u_1} = 0, \ldots, \frac{\partial F}{\partial u_m} = 0$$

which are same as equations (7) and (6).

Thus, the stationary points of f may be found by determining the stationary points of the function F, where

$$F = f + \lambda_1\phi_1 + \lambda_2\phi_2 + \ldots + \lambda_m\phi_m$$

and considering all the variables as independent variables.

A stationary point will be an extreme point of f if d^2F keeps the same sign, and will be a maxima or minima according as d^2F is negative or positive.

Note: It may be easier to deal with d^2F by expressing it in terms of two variables only or by expressing $dx\,dy$, $dy\,dz$, $dz\,dx$ in terms of dx^2, dy^2, dz^2, with the help of constraint conditions. Solved examples will illustrate the procedure.

Example 9. Find the shortest distance from the origin to the hyperbola

$$x^2 + 8xy + 7y^2 = 225, z = 0.$$

■ We have to find the minimum value of $x^2 + y^2$ (the square of the distance from the origin to any point in the xy plane) subject to the constraint

$$x^2 + 8xy + 7y^2 = 225$$

Consider the function

$$F = x^2 + y^2 + \lambda(x^2 + 8xy + 7y^2 - 225)$$

where x, y are independent variables and λ, a constant.

$$dF = (2x + 2x\lambda + 8y\lambda)dx + (2y + 8x\lambda + 14y\lambda)dy$$

$$\therefore \qquad \left.\begin{aligned}(1+\lambda)x + 4\lambda y = 0\\ 4\lambda x + (1+7\lambda)y = 0\end{aligned}\right\} \therefore\ \lambda = 1, -\frac{1}{9}$$

For $\lambda = 1$, $x = -2y$ and substitution in $x^2 + 8xy + 7y^2 = 225$, gives $y^2 = -45$, for which no real solution exists.

For $\lambda = -\frac{1}{9}$, $y = 2x$ and substitution in $x^2 + 8xy + 7y^2 = 225$, gives $x^2 = 5$, $y^2 = 20$, and so $x^2 + y^2 = 25$.

$$d^2F = 2(1+\lambda)dx^2 + 16\lambda\, dx\, dy + 2(1+7\lambda)dy^2$$

$$= \frac{16}{9}dx^2 - \frac{16}{9}dx\,dy + \frac{4}{9}dy^2, \text{ at } \lambda = -\frac{1}{9}$$

$$= \frac{4}{9}(2dx - dy)^2$$

$$> 0, \text{ and cannot vanish because } (dx, dy) \neq (0, 0).$$

Hence, the function $x^2 + y^2$ has a minimum value 25.

Note: Here F is a function of two variables and so its maximum or minimum values can be verified by the method of functions of two variables, $AC - B^2 > 0$ also.

Example 10. Find the maximum and minimum values of $x^2 + y^2 + z^2$ subject to the conditions $\frac{x^2}{4} + \frac{y^2}{5} + \frac{z^2}{25} = 1$, and $z = x + y$.

■ Let us consider a function F of independent variables x, y, z where

$$F = x^2 + y^2 + z^2 + \lambda_1\left(\frac{x^2}{4} + \frac{y^2}{5} + \frac{z^2}{25} - 1\right) + \lambda_2(x + y - z)$$

$$\therefore \quad dF = \left(2x + \frac{x}{2}\lambda_1 + \lambda_2\right)dx + \left(2y + \frac{2y}{5}\lambda_1 + \lambda_2\right)dy + \left(2z + \frac{2z}{25}\lambda_1 - \lambda_2\right)dz$$

As x, y, z are independent variables, we get

$$2x + \frac{x}{2}\lambda_1 + \lambda_2 = 0$$

$$2y + \frac{2y}{5}\lambda_1 + \lambda_2 = 0$$

$$2z + \frac{2z}{25}\lambda_1 - \lambda_2 = 0$$

$$\therefore \quad x = \frac{-2\lambda_2}{\lambda_1 + 4}, \quad y = \frac{-5\lambda_2}{2\lambda_1 + 10}, \quad z = \frac{25\lambda_2}{2\lambda_2 + 50}$$

Substituting in $x + y = z$, we get

$$\frac{2}{\lambda_1 + 4} + \frac{5}{2\lambda_1 + 10} + \frac{25}{2\lambda_1 + 50} = 0, \lambda_2 \neq 0 \qquad ...(1)$$

for if, $\lambda_2 = 0$, $x = y = z = 0$, but (0, 0, 0) does not satisfy the other condition of constraint.

Hence from (1), $17\lambda_1^2 + 245\lambda_1 + 750 = 0$, so that $\lambda_1 = -10, -75/17$.

For $\lambda_1 = -10$,

$$x = \frac{1}{3}\lambda_2, \; y = \frac{1}{2}\lambda_2, \; z = \frac{5}{6}\lambda_2$$

Substituting in $\frac{x^2}{4} + \frac{y^2}{5} + \frac{z^2}{25} = 1$, we get

$$\lambda_2^2 = 180/19 \text{ or } \lambda_2 = \pm 6\sqrt{5/19}$$

The corresponding stationary points are

$$(2\sqrt{5/19}, 3\sqrt{5/19}, 5\sqrt{5/19}), (-2\sqrt{5/19}, -3\sqrt{5/19}, -5\sqrt{5/19})$$

The value of $x^2 + y^2 + z^2$ corresponding to these point is 10.

For $\lambda_1 = -75/17$,

$$x = \frac{34}{7}\lambda_2, \; y = -\frac{17}{4}\lambda_2, \; z = \frac{17}{28}\lambda_2,$$

which on substitution in $\frac{x^2}{4} + \frac{y^2}{5} + \frac{z^2}{25} = 1$ give

$$\lambda_2 = \pm 140/(17\sqrt{646})$$

The corresponding stationary points are

$$(40/\sqrt{646}, -35/\sqrt{646}, 5/\sqrt{646}), (-40/\sqrt{646}, 35/\sqrt{646}, -5/\sqrt{646})$$

The value of $x^2 + y^2 + z^2$ corresponding to these points is 75/17.

Thus, the maximum value is 10 and the minimum 75/17.

Notes:

1. We have not theoretically established the existence of maximum or minimum value. We have simply shown that of all the possible values, 10 is the maximum and 75/17 the minimum.
2. Using constraint conditions, $dz = dx + dy; \frac{x}{4}dx + \frac{y}{5}dy + \frac{z}{25}dz = 0$, so that dz, dy and consequently d^2F may be expressed in terms of dx (or dx^2) alone. It can, then, be easily verified that 10 is a maximum value and 75/17 the minimum.

Example 11. Prove that the volume of the greatest rectangular parallelopiped, that can be inscribed in the ellipsoid $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$, is $\frac{8abc}{3\sqrt{3}}$.

■ We have to find the greatest value of $8xyz$ subject to the conditions

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1, \; x > 0, \; y > 0, \; z > 0 \qquad ...(1)$$

Let us consider a function F of three independent variables x, y, z, where

$$F = 8xyz + \lambda\left(\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} - 1\right)$$

$$\therefore \qquad dF = \left(8yz + \frac{2x\lambda}{a^2}\right)dx + \left(8zx + \frac{2y\lambda}{b^2}\right)dy + \left(8xy + \frac{2z\lambda}{c^2}\right)dz$$

At stationary points,

$$8yz + \frac{2x\lambda}{a^2} = 0,\ 8zx + \frac{2y\lambda}{b^2} = 0,\ 8xy + \frac{2z\lambda}{c^2} = 0 \quad ...(2)$$

Multiplying by x, y, z respectively and adding,

$$24xyz + 2\lambda = 0 \text{ or } \lambda = -12xyz \quad \text{[using (1)]}$$

Hence from (2), $x = a/\sqrt{3}$, $y = b/\sqrt{3}$, $z = c/\sqrt{3}$, and so

$$\lambda = -4abc/\sqrt{3}$$

Again

$$d^2F = 2\lambda\left(\frac{dx^2}{a^2} + \frac{dy^2}{b^2} + \frac{dz^2}{c^2}\right) + 16z\,dx\,dy + 16x\,dy\,dz + 16y\,dz\,dx$$

$$= -\frac{8abc}{\sqrt{3}}\,\Sigma\frac{1}{a^2}dx^2 + \frac{16}{\sqrt{3}}\Sigma c\,dx\,dy \quad ...(3)$$

Now from equations. (1), we have

$$x\frac{dx}{a^2} + y\frac{dy}{b^2} + z\frac{dz}{c^2} = 0 \text{ or } \frac{dx}{a} + \frac{dy}{b} + \frac{dz}{c} = 0 \quad ...(4)$$

Hence squaring,

$$\Sigma\frac{dx^2}{a^2} + 2\Sigma\frac{dx\,dy}{ab} = 0$$

or

$$abc\ \Sigma\frac{dx^2}{a^2} = -2\Sigma c\,dx\,dy$$

$$\therefore \qquad d^2F = -\frac{16}{\sqrt{3}}\,abc\ \Sigma\frac{dx^2}{a^2}$$

which is always negative.

Hence $\left(\frac{a}{\sqrt{3}}, \frac{b}{\sqrt{3}}, \frac{c}{\sqrt{3}}\right)$ is a point of maxima and the maximum value of $8xyz$ is $\frac{8abc}{3\sqrt{3}}$.

Note: The sign of d^2F can also be decided by expressing it in terms of dx and dy alone, by putting into (3) the value of dz from (4).

Example 12. Show that the length of the axes of the section of the ellipsoid $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$ by the plane $lx + my + nz = 0$ are the roots of the quadratic in r^2,

$$\frac{l^2a^2}{r^2 - a^2} + \frac{m^2b^2}{r^2 - b^2} + \frac{n^2c^2}{r^2 - c^2} = 0\,.$$

■ We have to find the stationary values of the function r^2 where $r^2 = x^2 + y^2 + z^2$, subject to the two equations of condition

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1 \qquad ...(1)$$

$$lx + my + nz = 0 \qquad ...(2)$$

Let us consider a function F of independent variables x, y, z,

$$F = x^2 + y^2 + z^2 + \lambda_1\left(\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} - 1\right) + 2\lambda_2(lx + my + nz)$$

$$\therefore \quad dF = 2\left(x + \frac{x\lambda_1}{a^2} + \lambda_2 l\right)dx + 2\left(y + \frac{y\lambda_1}{b^2} + \lambda_2 m\right)dy + 2\left(z + \frac{z\lambda_1}{c^2} + \lambda_2 n\right)dz$$

At stationary points,

$$x + \frac{x}{a^2}\lambda_1 + l\lambda_2 = 0,\ y + \frac{y}{b^2}\lambda_1 + m\lambda_2 = 0,\ z + \frac{z}{c^2}\lambda_1 + n\lambda_2 = 0 \qquad ...(3)$$

Multiplying by x, y, z, respectively and adding, we get

$$\lambda_1 = -(x^2 + y^2 + z^2) = -r^2$$

$$\therefore \quad x = \frac{a^2 l\lambda_2}{r^2 - a^2},\ y = \frac{b^2 m\lambda_2}{r^2 - b^2},\ z = \frac{c^2 n\lambda_2}{r^2 - c^2}$$

But $$0 = lx + my + nz = \lambda_2\left\{\frac{a^2l^2}{r^2 - a^2} + \frac{b^2m^2}{r^2 - b^2} + \frac{c^2n^2}{r^2 - c^2}\right\}$$

and since $\lambda_2 \neq 0$, we get the quadratic in r^2 giving the stationary values:

$$\frac{a^2l^2}{r^2 - a^2} + \frac{b^2m^2}{r^2 - b^2} + \frac{c^2n^2}{r^2 - c^2} = 0$$

Example 13. If the variables x, y, z satisfy the equation

$$\phi(x)\phi(y)\phi(z) = k^3 \qquad ...(1)$$

and $\phi(a) = k \neq 0,\ \phi'(a) \neq 0$, show that the function

$$f(x) + f(y) + f(z) \qquad ...(2)$$

has a maximum, when $x = y = z = a$, provided that

$$f'(a)\left\{\frac{\phi''(a)}{\phi'(a)} - \frac{\phi'(a)}{\phi(a)}\right\} > f''(a)$$

■ Let us consider a function

$$F = f(x) + f(y) + f(z) + \lambda\{\phi(x)\phi(y)\phi(z) - k^3\}$$

$$\therefore \quad dF = \Sigma\{f'(x) + \lambda\phi'(x)\phi(y)\phi(z)\}\,dx$$

At stationary points, $$\left.\begin{array}{l} f'(x) + \lambda\phi'(x)\phi(y)\phi(z) = 0 \\ f'(y) + \lambda\phi'(y)\phi(z)\phi(x) = 0 \\ f'(z) + \lambda\phi'(z)\phi(x)\phi(y) = 0 \end{array}\right\} \quad \ldots(3)$$

If the function has a maximum at (a, a, a), we must have

$$f'(a) + \lambda\phi'(a)\phi(a)\phi(a) = 0$$

or $$\lambda = -\frac{f'(a)}{\phi'(a)\phi^2(a)} = -\frac{f'(a)}{k^2\phi'(a)};\ \phi'(a) \neq 0,\ \phi(a) \neq 0$$

Again $$d^2F = \Sigma\{f''(x) + \lambda\phi''(x)\phi(y)\phi(z)\}dx^2 + 2\lambda\Sigma\phi'(x)\phi'(y)\phi(z)dx\,dy$$

At the stationary point (a, a, a),

$$d^2F = \{f''(a) + \lambda k^2\phi''(a)\}\,\Sigma dx^2 + 2\lambda k[\phi'(a)]^2\Sigma\, dx\,dy$$

From the given condition (1), we have

$$\Sigma\phi'(x)\phi(y)\phi(z)\, dx = 0$$

$\therefore$ $$k^2\phi'(a)(dx + dy + dz) = 0,\ \text{at } (a, a, a)$$

or

$$dx + dy + dz = 0$$

$\therefore$ $$2\Sigma\, dx\,dy = -\Sigma\, dx^2$$

$\therefore$ $$d^2F = \{f''(a) + \lambda k^2\phi''(a)\}\,\Sigma\, dx^2 - \lambda k[\phi'(a)]^2\,\Sigma dx^2$$

$$= \left\{f''(a) - f'(a)\frac{\phi''(a)}{\phi'(a)} + f'(a)\frac{\phi'(a)}{\phi(a)}\right\}\Sigma\, dx^2$$

For a maximum value at (a, a, a), d^2F is to be negative, *i.e.*,

$$f'(a)\left\{\frac{\phi''(a)}{\phi'(a)} - \frac{\phi'(a)}{\phi(a)}\right\} > f''(a).$$

Example 14. If $f(x, y, z) = (a^2x^2 + b^2y^2 + c^2z^2)/x^2y^2z^2$, where $ax^2 + by^2 + cz^2 = 1$, and a, b, c are positive, show that the minimum value of $f(x, y, z)$ is given by

$$x^2 = \frac{u}{2a(u + a)},\ y^2 = \frac{u}{2b(u + b)},\ z^2 = \frac{u}{2c(u + c)},$$

where u is the positive root of the equation

$$u^3 - (bc + ca + ab)\,u - 2abc = 0 \text{ (*Schlömilch*)}$$

■ Consider a function F of independent variables x, y, z, where

$$F = (a^2x^2 + b^2y^2 + c^2z^2)/x^2y^2z^2 + \lambda\,(ax^2 + by^2 + cz^2 - 1)$$

$$\therefore \qquad dF = \Sigma\left(2ax\lambda - \frac{2(b^2y^2+c^2z^2)}{x^3y^2z^2}\right)dx$$

At stationary points

$$\left.\begin{aligned} 2ax\lambda - \frac{2(b^2y^2+c^2z^2)}{x^3y^2z^2} &= 0 \\ 2by\lambda - \frac{2(c^2z^2+a^2x^2)}{x^2y^3z^2} &= 0 \\ 2cz\lambda - \frac{2(a^2x^2+b^2y^2)}{x^2y^2z^3} &= 0 \end{aligned}\right\}$$

or

$$ax^2\lambda = \frac{b^2y^2+c^2z^2}{x^2y^2z^2},\ by^2\lambda = \frac{c^2z^2+a^2x^2}{x^2y^2z^2},\ cz^2\lambda = \frac{a^2x^2+b^2y^2}{x^2y^2z^2} \qquad ...(1)$$

Adding, $$\lambda = \frac{2\Sigma a^2x^2}{x^2y^2z^2} \qquad (\because\ \Sigma ax^2 = 1)$$

Clearly λ is positive,

Again from (1),

$$\frac{b^2y^2+c^2z^2}{2ax^2} = \frac{c^2z^2+a^2x^2}{2by^2} = \frac{a^2x^2+b^2y^2}{2cz^2} = \frac{\Sigma a^2x^2}{1}$$

or $$\frac{a^2x^2}{1-2ax^2} = \frac{b^2y^2}{1-2by^2} = \frac{c^2z^2}{1-2cz^2} = \frac{\Sigma a^2x^2}{1} = \frac{u}{2}, \text{ say}$$

Clearly $u > 0$ and the coordinates of the stationary point are given by

$$\left.\begin{aligned} &2a^2x^2 = u(1-2ax^2) \Rightarrow x^2 = \frac{u}{2a(u+a)} \\ &\text{Similarly,} \quad y^2 = \frac{u}{2b(u+b)}, \qquad z^2 = \frac{u}{2c(u+c)} \end{aligned}\right\} \qquad ...(2)$$

Again substituting these values in the constraint condition, $\Sigma ax^2 = 1$, we get

$$1 = \frac{u}{2(u+a)} + \frac{u}{2(u+b)} + \frac{u}{2(u+c)}$$

which on simplification shows that u is a positive root of the equation,

$$u^3 - (bc+ca+ab)\,u - 2abc = 0 \qquad ...(3)$$

Again $$d^2F = \Sigma 2\left(a\lambda + \frac{3(b^2y^2+c^2z^2)}{x^4y^2z^2}\right)dx^2 + 2\Sigma\frac{4c^2}{x^3y^3}dx\,dy$$

Clearly $$F_{xx} = 2\left(a\lambda + \frac{3(b^2y^2 + c^2z^2)}{x^4y^2z^2}\right) > 0$$

Also

$$\begin{vmatrix} F_{xx} & F_{xy} \\ F_{xy} & F_{yy} \end{vmatrix} = F_{xx}F_{yy} - (F_{xy})^2 = 4\left(a\lambda + \frac{3(b^2y^2 + c^2z^2)}{x^4y^2z^2}\right)\left(b\lambda + \frac{3(c^2z^2 + z^2x^2)}{x^2y^4z^2}\right) - \left(\frac{4c^2}{x^3y^3}\right)^2$$

Comparison of the term containing $\frac{1}{x^3y^3}$ shows that this expression is positive.

It may be similarly shown that

$$\begin{vmatrix} F_{xx} & F_{xy} & F_{zx} \\ F_{xy} & F_{yy} & F_{yz} \\ F_{zx} & F_{yz} & F_{zz} \end{vmatrix} \text{ is also positive}$$

So all the three principal minors are positive.

Thus, d^2F is positive and so the function has a minimum value at the stationary point given by (2).

EXERCISE

1. Show that

   (*i*) if $2x + 3y + 4z = a$, the maximum value of $x^2y^3z^4$ is $\left(\frac{a}{9}\right)^9$.

   (*ii*) if $a^2x^2 + 2by^3 + z^4 = c^4$, the maximum value of x^4yz^2 is given by

   $$17a^2x^2 = 12c^4,\ 17by^3 = c^4,\ 17z^4 = 3c^4$$

2. If $xyz = abc$, the minimum value of $bcx + cay + abz$ is $3abc$.
3. If $x^2/a^2 + y^2/b^2 + z^2/c^2 = 1$, the maximum value of xyz is $abc/3\sqrt{3}$.
4. If $xyz = a^2(x + y + z)$, the minimum value of $yz + zx + xy$ is $9a^2$.
5. If $x^2 + y^2 = 1$, the minimum value of $(ax^2 + by^2)/(a^2x^2 + b^2y^2)^{1/2}$ is $2(ab)^{1/2}/(a + b)$.
6. If $xyz = k^3$, the product $(x + a)(y + b)(z + c)$ is a minimum, when $\frac{x}{a} = \frac{y}{b} = \frac{z}{c} = \frac{k}{(abc)^{1/3}}$; a, b, c are positive.
7. Show that the points on the ellipse $5x^2 - 6xy + 5y^2 = 4$ for which the tangent is at the greatest distance from the origin are $(1, 1)$ and $(-1, -1)$.
8. Show that the point on the sphere $x^2 + y^2 + z^2 = 1$ which is farther from $(2, 1, 3)$ is $(-2/\sqrt{14}, -1/\sqrt{14}, -3/\sqrt{14})$.
9. Show that the shortest distance from the origin to the curve of intersection of the surfaces $xyz = a$ and $y = bx$, where $a > 0, b > 0$, is $\sqrt[3]{a(b^2 + 1)/2b}$.
10. If $ax^2 + by^2 = ab$, show that the maximum and minimum values of $x^2 + xy + y^2$ will be the values of λ, given by the equation

    $$4(\lambda - a)(\lambda - b) - ab = 0$$

11. If $1/x + 1/y + 1/z = 1$, show that a stationary value of $a^3x^2 + b^3y^2 + c^3z^2$ is given by $ax = by = cz$, and this gives an extreme value if $abc(a+b+c)$ is positive. (To see the sign of d^2F, change it to two variables only with the help of the condition $\Sigma 1/x = 1$).

12. Find the shortest distance between the points P and Q, when P moves on the plane $x + y + z = 2a$, and Q on the ellipsoid $x^2/a^2 + y^2/b^2 + z^2/c^2 = 1$.

13. Find the maximum and the minimum distance from the origin to the curve of intersection of surface $(x^2 + y^2 + z^2)^2 = a^2x^2 + b^2y^2 + c^2z^2$ and the plane $lx + my + nz = 0$.

14. If $lx + my + nz = 1$, l, m, n are positive constants, show that the stationary value of $xy + yz + zx$ is

$$(2lm + 2mn + 2nl - l^2 - m^2 - n^2)^{-1}$$

and that this value is a maximum when it is positive.

15. Show that the function $xy + yz + zx$ has no extreme value, but has a maximum value when x, y, z are constrained by the condition $ax + by + cz = 1$, where a, b, c are positive constants satisfying the condition

$$2(ab + bc + ca) > (a^2 + b^2 + c^2)$$

16. If $\phi(x, y) = 0$, show that the determinant

$$\begin{vmatrix} f_{xx} + \lambda\phi_{xx} & f_{xy} + \lambda\phi_{xy} & \phi_x \\ f_{xy} + \lambda\phi_{xy} & f_{yy} + \lambda\phi_{yy} & \phi_y \\ \phi_x & \phi_y & 0 \end{vmatrix}$$

where λ is Lagrange's multiplier, is positive, in case the function $f(x, y)$ attains a maximum.

17. If $ax^2 + by^2 + cz^2 + 2fyz + 2gxz + 2hxy$ is equal to a constant k, and $lx + my + nz = 0$, find the maximum and minimum values of $x^2 + y^2 + z^2$.

18. If $ax + by + cz = 1$, show that in general $x^3 + y^3 + z^3 - 3xyz$ has the two stationary values 0 and $(a^3 + b^3 + c^3 - 3abc)^{-1}$, of which the first is a maximum or a minimum according as $a + b + c \gtrless 0$, but the second is not an extreme value. Discuss in particular, the cases when (*i*) $a + b + c = 0$, (*ii*) $a = b = c$.

19. If $x + y + z = 1$, find the stationary values of

$$x^3 + y^3 + z^3 + 3mxyz, \ (m \neq 2)$$

and show that the symmetrical stationary value is a maximum or minimum according as $m \lesseqgtr 2$, but the other stationary values are not extreme values.

Also show that $x^3 + y^3 + z^3 + 6xyz$ has only one stationary value and no extreme value.

17 Integration on R^2

(Line Integrals and Double Integrals)

In this chapter, we propose to consider two new concepts, the concepts of Line integrals and Double integrals in a two dimensional space. These integrals have important applications to geometry and physics.

Such problems as determining the mass of a material line from its density, computing the work of a field of force along a path and many others require the introduction of the so called *line integrals*, that is the integrals of bounded functions defined over curves. An ordinary single integral, $\int_a^b f\ dx$ is an integral of a function which is defined along a line segment (an interval of a coordinate axis). The corresponding kind of integral for a function which is defined along a curve, which might as well be called a *curvilinear integral*, is usually called a *line integral*, where line means, in general, a curved line. Likewise a *double integral* is an integral of a bounded function, defined on a bounded domain in the *xy*-plane.

1. LINE INTEGRALS

As was mentioned above, *line integrals* are integrals of functions, defined over curves. The curves may be plane curves or curves in space. In this section we shall study integrals over plane curves; those over space curves will be discussed later. Accordingly, we start with a study of the concept of curves.

1.1 Plane Curves (*Definitions*)

Intuitively we think of a curve as a one-dimensional configuration, like the path of a moving particle, or as something obtained by bending and twisting a straight line.

A *place curve* is a vector valued function $C = (X, Y)$ with domain as a subset of R and range as a subset of R^2.

But very often we do not make a distinction between a function and its range and take the range of C as the curve C itself. Thus a *plane curve* or a curve in R^2 is a set of points (x, y) for which

$$x = X(t), y = Y(t);\ a \le t \le b \qquad ...(1)$$

where X and Y are functions with domain $[a, b]$.

If f be a function of x with domain $[a, b]$, then the set of points

$$\{(x, y) : y = f(x);\ x \in [a, b]\}$$

is also a *plane curve* as may be seen on setting $y = f(t)$, $x = t$.

Similarly a curve in R^3 or a curve in space, is a vector valued function $C = (X, Y, Z)$ with domain as a subset of R and range, a subset of R^3, or equivalently, the set of points

$$\{(x, y, z): x = X(t), y = Y(t), z = Z(t), t \in [a, b]\}$$

In familiar terminology, we call (1) as parametric representation of the curve.

A curve is said to be *continuous*, if the functions X, Y and Z are continuous.

A point (x_1, y_1) is said to lie on a plane curve C, if there is a t for which

$$x_1 = X(t), y_1 = Y(t); \ a \le t \le b$$

If a curve is defined on the interval $[a, b]$, then the points corresponding to a and b are called its *end points*. The curve is said to be *closed*, if its end points coincide, *i.e.*, if

$$X(a) = X(b); \ Y(a) = Y(b)$$

A point P is said to be a *multiple point* of a curve, if the curve passes through P more than once, *i.e.*, if there is more than one value of t which yields P. Thus (x, y) is a *double point* of a plane curve C, if for two distinct values t_1, t_2 of t,

$$x = X(t_1) = X(t_2), y = Y(t_1) = Y(t_2) \quad t_1, t_2 \in [a, b]$$

A continuous curve is said to be *simple*, if it has no multiple points. Thus a continuous closed curve is *simple*, if the only multiple points are the coincident end points. A simple continuous closed curve is generally called a *Jordan curve* or a *Jordan arc* (or simply an *arc*).

A curve is said to be *smooth*, if it has no multiple points and X', Y' exist and are continuous and do not vanish simultaneously on $[a, b]$. A curve is said to be *piecewise smooth* on $[a, b]$, if it is composed of a finite number of smooth arcs.

Remark: A smooth curve is a continuously differentiable curve. It has a continuously turning tangent, for, X' and Y' do not vanish simultaneously and their quotient determines the direction of the tangent.

For the purpose of integration, a curve is generally taken to be *oriented*. It is said to be oriented in one way as t varies from a to b and the other way as t varies from b to a. Thus if it is denoted by C, when oriented in one way, it is denoted as $-C$, when oriented in the other way.

However, in the case of a closed curve, the curve is oriented in such a way that the enclosed area is always on the left as one moves along the curve in the positive direction.

1.2 Line Integral (*Definition*)

Let

$$x = X(t), y = Y(t); \ a \le t \le b$$

be a curve C.

Let a bounded function f be defined at every point of C (*i.e.*, the domain of f contains the curve).

Let $P = \{a = t_0, t_1, t_2, \ldots, t_n = b\}$ be any partition of $[a, b]$, and ξ_i any point of Δt_i.

Form the sum

$$S(P, f, X) = \Sigma f(X(\xi_i), Y(\xi_i)) \, \Delta x_i$$

If as the norm $\mu(P) \to 0$, the sum $S(P, f, X)$ tends to a finite limit, which is independent of the choice of ξ_i, the limit is denoted by

$$\int_C f(x, y)\, dx \quad \text{or} \quad \int_C f(X, Y)\, dX$$

and is called the *line integral* of f along the curve C.

This means that there exists a number I such that to each $\varepsilon > 0$, there corresponds $\delta > 0$, such that, for every partition P of $[\alpha, \beta]$ with norm $\mu(P) < \delta$, and for every choice of the point ξ_i in $[t_{i-1}, t_i]$,

$$|\Sigma f(X(\xi_i), Y(\xi_i))\, \Delta x_i - I| < \varepsilon$$

Thus, we have

$$I = \int_C f(x, y)\, dx \qquad \text{...(2)}$$

$$= \int_C f(X, Y)\, dX \qquad \text{...(3)}$$

Notes:

1. Relation (3) depicts the line integral as a Riemann-Stieltjes integral but we shall give here an exposition that is independent of it. However the reader will do well to note the similarity.
2. If instead of f we take a vector-valued function $F = (f, g)$ defined on C and formed the sum

$$\Sigma\{f(X(\xi_i), Y(\xi_i))\, \Delta x_i + g(X(\xi_i), Y(\xi_i))\, \Delta y_i\}$$

we would get the line integral

$$\int_C \{f(X, Y)\, dx + g(X, Y)\, dy\}$$

or

$$\int_C (f\, dx + g\, dy)$$

which is discussed in the next section.

1.3 A Sufficient Condition for Existence

If X, Y and f are continuous, and X possesses a continuous derivative X' on $[\alpha, \beta]$, then the line integral of f along the curve C, $x = X(t)$, $y = Y(t)$ exists.

Let $P = \{a = t_0, t_1, t_2, \dots, t_n = b\}$ be any partition of $[a, b]$. Since X is derivable in Δt_i there exists $\eta_i \in \Delta t_i$ such that

$$\Delta x_1 = X(t_i) - X(t_{i-1}) = X'(\eta_i)\, \Delta t_i$$

$$\therefore \quad S(P, f, X) = \sum_i f(X(\xi_i), Y(\xi_i))\, \Delta x_i,\ \xi_i \in \Delta t_i$$

$$= \sum_i f(X(\xi_i), Y(\xi_i))\, X'(\xi_i)\, \Delta t_i + \sum_i f(X(\xi_i), Y(\xi_i))\, \{X'(\eta_i) - X'(\xi_i)\}\, \Delta t_i$$

$$= S_1 + S_2 \quad \text{(say)} \qquad \text{...(1)}$$

Since $f(X, Y)$ and X' are continuous, it follows (Th. 12, Ch. 9) that S_1 tends to a finite limit,

$$\int_a^b f(X, Y) X'\, dt, \text{ as the norm } \mu(P) \to 0.$$

Again, since $f(X, Y)$ is continuous, it is bounded and therefore a positive number K exists such that

$$|f(X, Y)| \le K, \quad \forall\ t \in [a, b]$$

Also since X' is continuous, it is integrable and accordingly there exists for any $\varepsilon > 0$, a positive number δ such that for every partition P with norm less than δ, we have

$$U(P, X') - L(P, X') = \sum(M_i - m_i)\,\Delta t_i < \varepsilon/K$$

where M_i, m_i are the bounds of X' in Δt_i.

Let P be such a partition.

$$\therefore \qquad |S_2| \le \sum |f(X(\xi_i), Y(\xi_i))| \cdot |X'(\eta_i) - X'(\xi_i)|\,\Delta t_i$$

$$\le K\sum(M_i - m_i)\,\Delta t_i < K(\varepsilon/K) = \varepsilon \qquad ...(2)$$

Thus, $S_2 \to 0$, as $\mu(P) \to 0$.

$$\Rightarrow \qquad \lim_{\mu(P)\to 0} S(P, f, X) \text{ exists and equals } \int_a^b f(X, Y)X'\,dt$$

Hence, the line integral of f along the curve C exists, and we have the equality

$$\int_C f(x, y)\,dx = \int_a^b f(X, Y)X'\,dt \qquad ...(3)$$

Notes:

1. Inequality (2) may be deduced by the following consideration as well:
   Since X' is continuous, for $\varepsilon > 0$ there exist $\delta > 0$ such that
   $$|X'(\eta_i) - X'(\xi_i)| < \varepsilon/K(b-a),$$
   whenever $|\eta_i - \xi_i| < \delta$. Take P to be a partition with norm less than δ.
2. The line integral reduces to ordinary integral when the path of integration (the curve C) is an interval on the x or y-axis.

Remarks:

1. In a similar way, we define and examine the existence of the line integrals
   $$\int_C g(x, y)\,dy \text{ and } \int_C \{f(x, y)\,dx + g(x, y)\,dy\}$$
2. It may be easily seen that the line integral $\int_C f(x, y)\,dy$ along the curve C,
   $$y = \phi(x),\ a \le x \le b$$
   is equal to ordinary integral $\int_a^b f(x, \phi(x))\,\phi'(x)\,dx$.

1.4 Properties of Line Integrals

Since a line integral can be reduced (Relation 3 § 1.3) to an ordinary definite integral, the basic properties of the integral are almost completely analogous to those of the definite integral.

(*i*) If k is a constant,

$$\int_C kf(x, y)\,dx = k\int_C f(x, y)\,dx$$

(*ii*) If f and g are integrable on C, then $f \pm g$ are also integrable on C and

$$\int_C (f \pm g)\, dx = \int_C f\, dx \pm \int_C g\, dx$$

(*iii*) If f is a non-negative integrable function, then

$$\int_C f(x, y)\, dx \geq 0$$

(*iv*) If f is integrable on C, then $|f|$ is also integrable on C, and

$$\left| \int_C f(x, y)\, dx \right| \leq \int_C |f(x, y)|\, dx$$

(*v*) If an arc AB is composed of two arcs AC and CB, then

$$\int_{AB} f(x, y)\, dx = \int_{AC} f(x, y)\, dx + \int_{CB} f(x, y)\, dx$$

which may be easily extended to a finite number of arcs.

Example 1. Evaluate the integral $\int_C (x^2\, dx + xy\, dy)$ taken along: (*i*) the line segment from (1, 0) to (0, 1), (*ii*) the quarter circle $x = \cos t$, $y = \sin t$, joining the same points.

■ (*i*) The equation of the line joining (1, 0) and (0, 1) is $x + y = 1$.

$$\therefore \int_C (x^2\, dx + xy\, dy) = \int_1^0 \{x^2 - x(1 - x)\}\, dx = \int_1^0 (2x^2 - x)\, dx = -1/6$$

(*ii*) $\int_C (x^2\, dx + xy\, dy) = \int_0^{\pi/2} (-\cos^2 t \sin t + \cos^2 t \sin t)\, dt = 0.$

Example 2. Compute the integral $\int_C xy\, dx$ along the arc of the parabola $x = y^2$ from (1, –1) to (1, 1).

■ *First method.* When the equation of the curve of integration is taken by expressing x as a single valued function of y, we put $x = y^2$ so that

$$= \int_C xy\, dx = \int_{-1}^{1} y^2 \cdot y \cdot 2y\, dy = 4/5$$

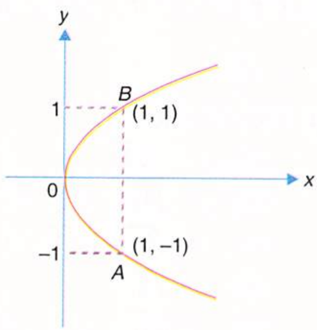


Fig. 1

Second Method. If y is expressed as a function of x, the arc AB has to be split into two parts, AO and OB, whose equations are

$$y = -\sqrt{x},\ y = +\sqrt{x}$$

$$\therefore \qquad \int_C xy\,dx = \int_{AO} xy\,dx + \int_{OB} xy\,dx = \int_1^0 -x\sqrt{x}\,dx + \int_0^1 x\sqrt{x}\,dx = 4/5$$

Example 3. Find the value of

$$\int_C \{(x + y^2)\,dx + (x^2 - y)\,dy\}$$

taken in the clockwise sense along the closed curve C formed by $y^3 = x^2$ and the chord joining (0, 0) and (1, 1).

■ The curve C consists of the arc OA, $(y^3 = x^2)$ and the chord AO, $(y = x)$

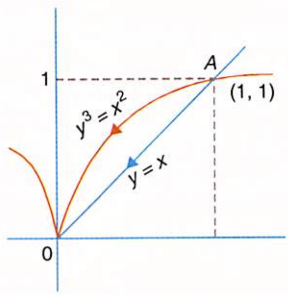


Fig. 2

$$\therefore \qquad \int_C \{(x + y^2)\,dx + (x^2 - y)\,dy\}$$

$$= \int_0^1 \{(x + x^{4/3})\,dx + (y^3 - y)\,dy\} + \int_1^0 \{(x + x^2)\,dx + (x^2 - x)\,dx\}$$

$$= \int_0^1 (x^{4/3}\,dx + y^3\,dy) + \int_1^0 2x^2\,dx = 1/84$$

Example 4. Find the value of $\int_C (x^2 y\,dx + xy^2\,dy)$ taken in the clockwise sense along the hexagon whose vertices are $(\pm 3a, 0)\ (\pm 2a, \pm\sqrt{3a})$.

■ Equations of the lines forming the curve C are

$$AB,\ y = -\sqrt{3a}$$
$$BC,\ y - \sqrt{3x} + 3\sqrt{3a} = 0$$
$$CD,\ y + \sqrt{3x} - 3\sqrt{3a} = 0$$
$$DE,\ y = \sqrt{3a}$$
$$EF,\ y - \sqrt{3x} - 3\sqrt{3a} = 0$$

$$FA, \ y + \sqrt{3x} + 3\sqrt{3a} = 0$$

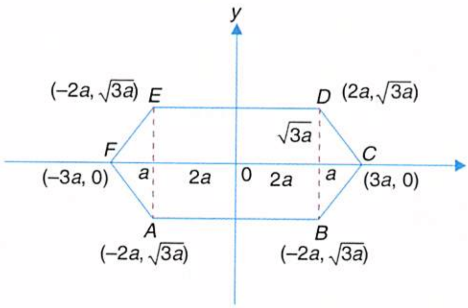


Fig. 3

Let

$$I_C = \int_C (x^2 y \, dx + xy^2 \, dy)$$

$\therefore$

$$I_C = I_{AF} + I_{FE} + I_{ED} + I_{DC} + I_{CB} + I_{BA}$$

$$= \left\{ -\int_{-2a}^{-3a} x^2\left(\sqrt{3x} + 3\sqrt{3a}\right) dx - \int_{-\sqrt{3a}}^{0} y^2\left(\frac{y}{3a} + 3a\right) dy \right\}$$

$$+ \left\{ \int_{-3a}^{-2a} x^2\left(\sqrt{3x} + 3\sqrt{3a}\right) dx + \int_{0}^{\sqrt{3a}} y^2\left(\frac{y}{\sqrt{3}} - 3a\right) dy \right\}$$

$$+ \int_{-2a}^{2a} \sqrt{3a}x^2 \, dx + \left\{ \int_{2a}^{3a} x^2\left(-\sqrt{3x} + 3\sqrt{3a}\right) dx \right.$$

$$\left. + \int_{\sqrt{3a}}^{0} y^2\left(\frac{-y}{\sqrt{3}} + 3a\right) dy \right\} + \left\{ \int_{3a}^{2a} x^2\left(\sqrt{3x} - 3\sqrt{3a}\right) dx \right.$$

$$\left. + \int_{0}^{-\sqrt{3a}} y^2\left(\frac{y}{\sqrt{3}} + 3a\right) dy \right\} - \left\{ \int_{3a}^{-2a} \sqrt{3a}x^2 \, dx \right.$$

$$= 4\int_{2a}^{3a} - x^2\left(\sqrt{3x} - 3\sqrt{3a}\right) dx$$

$$+ 4\int_{0}^{\sqrt{3a}} y^2\left(\frac{y}{\sqrt{3}} - 3a\right) dy + 2\sqrt{3a}\int_{-2a}^{2a} x^2 \, dx$$

$$= 4a^4\left(\frac{65\sqrt{3}}{4} - 19\sqrt{3}\right) + 4a^4\left(\frac{9}{4\sqrt{3}} - 3\sqrt{3}\right) + \frac{32}{\sqrt{3}}a^4 = \frac{38}{\sqrt{3}}a^4.$$

EXERCISE

1. Show that

$$\int_C \{(x-y)^3\, dx + (x-y)^3\, dy\} = 3\pi a^4$$

taken along the circle $x^2 + y^2 = a^2$ in the counter-clockwise direction.

2. Show that

$$\int_C \frac{y^2\, dx - x^2\, dy}{x^2 + y^2} = -\frac{4a}{3}$$

where C is the semicircle $x = a \cos t$, $y = a \sin t$ from $t = 0$ to $t = \pi$.

3. Compute $\int_\Gamma \frac{dx}{x+y}$, where Γ is the curve $x = at^2$, $y = 2at$, $0 \le t \le 2$.

4. Evaluate $\int_C (y\, dx - x\, dy)$, where C is the ellipse $x = a \cos t$, $y = b \sin t$ taken in the clockwise direction.

5. Show that

$$\int_\Gamma (y\, dx - x\, dy) = 24\pi$$

where Γ is the arc of the cycloid $x = 2(t - \sin t)$, $y = 2(1 - \cos t)$, joining the points $(0, 0)$ and $(4\pi, 0)$. [*Hint*: t varies from 0 to 2π.]

6. Evaluate $\int_C \{(2a - y)\, dx - (a - y)\, dy\}$, where C is the arc of the cycloid (from the origin) $x = a(t - \sin t)$, $y = a(1 - \cos t)$.

7. Show that

$$\int_C \frac{x^2\, dy - y^2\, dx}{x^{5/3} + y^{5/3}} = \frac{3\pi}{16} a^{4/3}$$

where C is the quarter of the astroid $x = a \cos^3 t$, $y = a \sin^3 t$, from the point $(a, 0)$ to the point $(0, a)$.

8. Find the value of $\int_\Gamma (x^2 + y^2)\, dy$, taken in the counter-clockwise sense along the quadrilateral with vertices $(0, 0)$, $(2, 0)$, $(4, 4)$ and $(0, 4)$.

9. Compute

$$\int_C (x^2 y\, dx + y^2 x\, dy)$$

taken in the clockwise sense along the hexagon whose vertices are $(\pm 2a, 0)$, $(\pm a, \pm\sqrt{3}a)$.

10. Show that $\int_C (xy^2\, dy - x^2 y\, dx)$, taken in the counter-clockwise sense along the cardioid

$$r = a(1 + \cos\theta) \text{ is } 35a^4\pi/16.$$

ANSWERS

3. $2 \log 2$, 4. $2\pi ab$, 6. πa^2, 8. $112/3$, 9. 0.

2. DOUBLE INTEGRALS

In this section we shall discuss a double integral over a finite rectangle. We shall try to follow as closely as possible the analogy with the theory of single integrals developed in the earlier chapters. Let us first study the definitions of certain concepts which will be needed in the subsequent work.

2.1 Definitions

Domain. A domain is an open *connected* set of points, *i.e.*, an open set any two of whose points can be joined by a broken line having a finite number of segments, all of whose points belong to the set.

A domain is *bounded*, if all its points lie inside some rectangle (square).

Boundary. A point is called a *boundary point* of a set S, if every neighbourhood of it contains points of S as well as those not belonging to S. Thus a point is a *boundary point* of S, if it is a limit point of S as also of the complement of S. The set of all the boundary points of a set is called the *boundary* of the set.

Region. A *region* or a *closed domain* is a closed point set consisting of a bounded domain plus its boundary points.

We shall assume further that the boundary of a region consists of a finite number of closed curves, simple or Jordan curves, that do not cross themselves nor each other. Symbol E will, generally, be employed to denote a region, and symbol S to denote its area.

A region is *simply connected*, if its boundary consists of a single closed curve.

A closed domain (region) E is said to be *Regular* or *Quadratic* with respect to y-axis, if it is bounded by curve of the form:

$$y = \phi(x),\ y = \Psi\ (x);\ x = a,\ x = b$$

where ϕ and Ψ are continuous and $\phi(x) \geq \Psi\ (x), \quad \forall\ x \in [a, b]$

Thus a domain which is *regular* (*quadratic*) with respect to y-axis is such that a line parallel to y-axis and lying between $x = a$, $x = b$ meets the boundary of E in just two points. Domains, quadratic with respect to x-axis are defined in a similar fashion. The domain is said to be *piecewise regular* (*quadratic*) with respect to an axis, if it can be divided into a finite number of domains each of which is regular with respect to that axis.

A domain, regular with respect to all the axes is called a *regular* (*quadratic*) *domain*. A *piecewise regular domain* is defined similarly.

The contour of a closed region is said to be described in a *positive sense*, if the enclosed region always lies to the left as one advances along the contour. The area of the enclosed region is then taken to be positive.

Diameter. The diameter of a region E is the length of the largest line segment that joins two points of E.

It may be observed that if a region varies so that its diameter approaches zero, then its area also approaches zero. The converse, however, is not true.

A *partition* of a region E is the set of closed curves which divide E into a finite number of sub-regions E_i of area ΔS_i, $i = 1, 2, 3,\ldots$

The *norm* $\mu(P)$ of a partition P is the maximum diameter of the sub-regions produced by the partition P.

Partition of a rectangle. In the xy-plane, consider a rectangle bounded by the lines

$$x = a, x = b, y = c, y = d, a < b; c < d$$

Let the region within this rectangle including its boundary be denoted by $[a, b; c, d]$ or by R as convenient. Let us define its area to be the number $(b-a)(d-c)$.

Let $P_1 = \{a = x_0, x_1, \ldots, x_n = b\}$, and $P_2 = \{c = y_0, y_1, \ldots, y_m = d\}$ be partitions of the intervals $[a, b]$ and $[c, d]$ respectively. Lines drawn parallel to the axes through the points of the partitions P_1 and P_2, give rise to a partition P of the rectangle R into mn sub-rectangles.

$$[x_{i-1}, x_i\ y_{j-1}, y_j], i = 1, 2, \ldots, n, j = 1, 2, \ldots, m$$

Partition P is in fact the Cartesian product of P_1 and P_2, where

$$P = P_1 \times P_2 = \{(x_i, y_j); x_i \in P_1, y_j \in P_2\}$$

so that P consists of all the grid points.

We shall use the same symbol ΔR_{ij} to denote the sub-rectangle, $[x_{i-1}, x_i; y_{j-1}, y_j]$ as also its area

$$(x_i - x_{i-1})(y_j - y_{j-1}) = \Delta x_i \cdot \Delta y_j$$

If $\mu(P_1) = \Delta x_r$ and $\mu(P_2) = \Delta y_s$ be the norms of the partitions P_1 and P_2 respectively, then the sub-rectangle $\Delta R_{rs} = [x_{r-1}, x_r; y_{s-1}, y_s]$ is called the norm of the partition P, and is denoted by $\mu(P)$. Clearly the area of each sub-rectangle, ΔR_{ij} tends to zero as $\mu(P) \to 0$.

Refinement. A partition P^* is said to be a refinement of P if $P^* \supseteq P$, *i.e.*, every line of P is also a line of P^*.

2.2 Integration Over a Rectangle

Let f be a bounded function of x, y over a rectangle $R = [a, b; c, d]$. Let P be a partition of R which divides R into mn sub-rectangles $\Delta R_{ij}, i = 1, 2, \ldots, n; j = 1, 2, \ldots m$. Let M_{ij}, m_{ij} be the upper and the lower bounds of f in ΔR_{ij}. Consider the two sums

$$U(P, f) = \sum_i \sum_j M_{ij}\ \Delta R_{ij} = \sum_i \sum_j M_{ij}\ \Delta x_i\ \Delta y_j$$

$$L(P, f) = \sum_i \sum_j m_{ij} \cdot \Delta R_{ij} = \sum_i \sum_j m_{ij}\ \Delta x_i\ \Delta y_j$$

respectively, called the *upper* and the *lower* (Darboux) *sums*.

It may be easily shown that

$$m(b-a)(d-c) \le L(P, f) \le U(P, f) \le M(b-a)(d-c) \qquad \ldots(1)$$

where M, m are the upper and the lower bounds of f in R.

As in Riemann integrals, considering all partitions of the rectangle R, we find that the sets of the upper and the lower sums are bounded and possess the infimum and the supremum.

The infimum of the set of upper sums is called the *upper integral*, I^u, and the supremum of the set of lower sums is called the *lower integral*, I_l of f over R and denoted as

$$I^u = \overline{\iint_R} f\, dx\, dy, \quad \text{upper integral}$$

$$I_l = \underline{\iint_R} f\, dx\, dy, \quad \text{lower integral}$$

f is said to be integrable when the two integrals are equal and the common value I is called the Double integral of f over R, denoted by $\iint_R f\, dx\, dy$.

Thus

$$I = \overline{\iint_R} f\, dx\, dy = \underline{\iint_R} f\, dx\, dy = \iint_R f\, dx\, dy\,.$$

Remarks:

1. For all partitions P of R we see that

$$L(P, f) \le I \le U(P, f)$$

Thus the integral $\iint_R f\, dx\, dy$ can be thought of as a number I which lies between $L(P, f)$ and $U(P, f)$ for all partitions of the rectangle (*i.e.*, independent of the partitions).

2. It may be seen from (1) that

$$m(b-a)(d-c) \le \iint_R f\, dx\, dy \le M(b-a)(d-c)$$

which implies that

$$\iint_R f\, dx\, dy = \mu(b-a)(d-c), \text{ where } m \le \mu \le M$$

If f is continuous on R, then

$$\iint_R f\, dx\, dy = (b-a)(d-c)\, f(\xi, \eta), (\xi, \eta) \in R$$

3. If $f(x, y) = 1$, then

$$L(P, f) = \sum\sum \Delta R_{ij} = (b-a)(d-c) = U(P, f)$$

$\Rightarrow$

$$I_l = I^u = (b-a)(d-c)$$

i.e., the double integral exists and is equal to the area of the rectangle R.

Hence $\iint_R dx\, dy$ is the area of the rectangle R, $dx\, dy$ being the elementary area in cartesian coordinates.

We can easily establish the following properties of the Darboux sums and the upper and lower integrals. The proofs are on the same lines as for the single integrals.

1. If P^* is a refinement of a partition P, then for a bounded function f,

   (*i*) $L(P^*, f) \ge L(P, f)$

   (*ii*) $U(P^*, f) \le U(P, f)$

   (*iii*) for any two partitions P_1 and P_2

$$L(P_1, f) \le U(P_2, f)$$

(*iv*) $\underline{\iint}_R f\,dx\,dy \le \overline{\iint}_R f\,dx\,dy$

2. *Darboux theorem*. To every $\varepsilon > 0$, there corresponds $\delta > 0$ such that

$$U(P, f) < I^u + \varepsilon,\ \ L(P, f) > I_l - \varepsilon$$

for every partition P of R with norm $\mu(P) < \delta$.

2.3 Conditions of Integrability

I. *First form.* *A bounded function f is integrable over a rectangle R, if and only if to every* $\varepsilon > 0$, *there corresponds* $\delta > 0$, *such that for every partition P of R with norm* $\mu(P) < \delta$, *the oscillatory sum,* $U(P, f) - L(P, f) < \varepsilon$.

Necessary. The bounded function f being integrable

$$I^u = I_l = I$$

Let ε be any positive number. By Darboux theorem there exists $\delta > 0$ such that, for every partition P with norm $\mu(P) < \delta$, we have

$$U(P,f) < I^u + \frac{1}{2}\varepsilon = I + \frac{1}{2}\varepsilon$$

$$L(P,f) < I_l + \frac{1}{2}\varepsilon = I - \frac{1}{2}\varepsilon$$

which give $$U(P, f) - L(P, f) < \varepsilon$$

Sufficient. Let ε be any positive number.

For any partition P with norm $\mu(P) < \delta$, we are given that

$$U(P, f) - L(P, f) < \varepsilon$$

Also for any partition P,

$$L(P, f) \le I_l \le I^u \le U(P, f)$$

$\Rightarrow$ $$I^u - I_l \le U(P, f) - L(P, f) < \varepsilon$$

i.e., a non-negative number, $I^u = I_l$ is less than any arbitrary positive number ε, which is possible only when

$$I^u - I_l = 0 \ \text{ or } \ I^u = I_l$$

so that the function is integrable.

II. *Second form.* *A bounded function f is integrable on a rectangle R if and only if for every* $\varepsilon > 0$, *there corresponds a partition P of R, such that*

$$U(P, f) - L(P, f) < \varepsilon$$

The proof is similar to that for the functions of a single variable.

2.4 Integral as a Limit of Sums

Let f be a bounded function over a rectangle $R = [a, b; c, d]$. Let (ξ_{ij}, η_{ij}) be a point of the sub-rectangle ΔR_{ij} corresponding to any partition P of R.

Form the sum

$$S(P, f) = \sum_{i=1}^{n} \sum_{j=1}^{m} f(\xi_{ij}, \eta_{ij}) \Delta R_{ij}$$

If the limit of this sum, when $\mu(P) \to 0$ exists, for all partitions P of R and all positions of the point (ξ_{ij}, η_{ij}) in ΔR_{ij}, we say the function f is integrable over R, and

$$\lim_{\mu(P)\to 0} S(P, f) = \iint_R f\, dx\, dy$$

called the *double integral* of f over the rectangle R.

Remark: With the increase in their number, the size of the sub-rectangles decreases and so $\mu(P) \to 0$ as $mn \to \infty$. Thus, we may replace the condition $\mu(P) \to 0$ by $mn \to \infty$ in above definition, so that

$$\lim_{mn\to\infty} S(P, f) = \iint_R f\, dx\, dy$$

2.5 Some Integrable Functions

I. *Every continuous function is integrable.*

Let f be continuous on a rectangle $R = [a, b; c, d]$ and let ε be any positive number.

By the property of uniform continuity of f, a partition P of R exists such that the oscillation $(M_{ij} - m_{ij})$ of f in every subrectangle of P is less than $\varepsilon/(b-a)(d-c)$.

Hence for such a partition,

$$U(P, f) - L(P, f) = \Sigma\Sigma(M_{ij} - m_{ij})\Delta R_{ij} < [\varepsilon/(b-a)(d-c)]\, \Sigma\Sigma\, \Delta R_{ij} = \varepsilon$$

$\Rightarrow$ f is integrable over R.

II. *A bounded function f, having a finite number of points of discontinuity on a rectangle R is integrable on R.*

Let $\varepsilon > 0$ be any number, and M, m the bounds of f. Let us consider a partition of R such that all the points of discontinuity get enclosed in a finite number of sub-rectangles, the sum of whose areas is less than $\varepsilon/2(M-m)$. The oscillation of f being less than $(M-m)$, the part of the oscillatory sum $\{U(P, f) - L(P, f)\}$ arising from these rectangles or from the sub-rectangles into which they may be further partitioned, is less than $\frac{1}{2}\varepsilon$. The remaining sub-rectangles, where the function is continuous, may be further partitioned (as in I above) so that the part of oscillatory summarising from them is less than $\frac{1}{2}\varepsilon$.

Thus we have found a partition of R such that the corresponding oscillatory sum is less than any given positive number ε.

Hence, the function f is integrable on R.

III. *If a function f is bounded on $R = [a, b; c, d]$, and its points of discontinuity lie on a finite number of curves of the form $y = \phi(x)$ or $x = \Psi(y)$, etc., where $\phi, \Psi, \ldots$ are continuous, then f is integrable on R.*

Let the number of curves in question be p.

Let $\varepsilon > 0$ be any number.

Because of continuity of ϕ in $[a, b]$, there exists a number $\delta > 0$, such that for all partitions P, with $\mu(P) < \delta$, the oscillation of ϕ in every sub-interval is less than $\varepsilon/(b-a)$. Let $P = \{a = x_0, x_1, ..., x_n = b\}$ be such a partition.

The points of $y = \phi(x)$, corresponding to the values of $x \in \Delta x_i$, clearly belong to the rectangle $[x_{i-1}, x_i\,; m_i, M_i]$ of area $(x_i - x_{i-1})(M_i - m_i)$; M_i, m_i being the bounds of ϕ in Δx_i. Thus all points of the curve have been enclosed in rectangles of total area $\Sigma(x_i - x_{i-1})(M_i - m_i)$, which is less than ε/p.

Dealing with the other curves in a similar manner, we see that all the curves can be enclosed in a finite number of rectangles of total area less than any positive number ε. Now proceeding as in II above, we can prove the required result.

2.6 Some Theorems

Following are some of the simple theorems. Their proofs being straightforward, are left to the reader.

If the functions f, f_1, f_2 of x, y are integrable on a rectangle R, then

I. $\iint_R kf\, dx\, dy = k \iint_R f\, dx\, dy$, k is constant.

II. $\iint_R (f_1 + f_2)\, dx\, dy = \iint_R f_1\, dx\, dy + \iint_R f_2\, dx\, dy$.

III. The product $f_1 f_2$ is integrable over R.

IV. The quotient f_1/f_2 is integrable over R, if $|f_2| \geq k > 0$ on R.

V. When f is integrable on R so is $|f|$, and

$$\left|\iint_R f\, dx\, dy\right| \leq \iint_R |f|\, dx\, dy.$$

VI. If the rectangle R is subdivided into a finite number of sub-rectangles $R_1, R_2, \ldots, R_m$, then

$$\iint_R f\, dx\, dy = \iint_{R_1} f\, dx\, dy + \iint_{R_2} f\, dx\, dy + \ldots + \iint_{R_m} f\, dx\, dy.$$

VII. *Mean Value Theorem*. If f and g are integrable and g keeps the same sign on R, and m, M are the lower and upper bounds of f, then

(*i*) $m \iint_R g\, dx\, dy \leq \iint_R fg\, dx\, dy \leq M \iint_R g\, dx\, dy$, when $g(x, y) \geq 0$

$m \iint_R g\, dx\, dy \geq \iint_R fg\, dx\, dy \geq M \iint_R g\, dx\, dy$, when $g(x, y) \leq 0$

(*ii*) $\iint_R fg\, dx\, dy = \mu \iint_R g\, dx$, $m \leq \mu \leq M$

(*iii*) if f is continuous on R,

$$\iint_R fg\, dx\, dy = f(\xi, \eta) \iint_R g\, dx\, dy, (\xi, \eta) \in R.$$

Example 5(a). Evaluate $\int\int_R (x+y)\, dx\, dy$ on the rectangle $R = [a, b; c, d]$.

■ Let $P_1 = \{a = x_0, x_1, \ldots, x_n = b\}$ be an arbitrary partition of $[a, b]$ and

$P_2 = \{c = y_0, y_1, \ldots, y_m = d\}$ that of $[c, d]$. Then

$P = P_1 \times P_2 = \{(x_i, y_j) : x_i \in P_1, y_j \in P_2\}$

is an arbitrary partition of R.

On each sub-rectangle, $\Delta R_{ij} = [x_{i-1}, x_i ; y_{j-1}, y_j]$, $(x_{i-1} + y_{j-1})$ and $(x_i + y_j)$ are the lower and the upper bounds of $x + y$. Thus

$$L(P, f) = \sum_i \sum_j (x_{i-1} + y_{j-1})\, \Delta x_i\, \Delta y_j$$

$$U(P, f) = \sum_i \sum_j (x_i + y_j)\, \Delta x_i\, \Delta y_j$$

Also

$$x_{i-1} + y_{j-1} \le \frac{1}{2}(x_i + x_{i-1}) + \frac{1}{2}(y_j + y_{j-1}) \le x_i + y_j$$

$\therefore$
$$L(P,f) \le \sum_i \sum_j \left\{\frac{1}{2}(x_i + x_{i-1}) + \frac{1}{2}(y_j + y_{j-1})\right\} \Delta x_i\, \Delta y_j \le U(P,f)$$

But

$$\sum_i \sum_j \left\{\frac{1}{2}(x_i + x_{i-1}) + \frac{1}{2}(y_j + y_{j-1})\right\} \Delta x_i\, \Delta y_j$$

$$= \frac{1}{2}\sum_i \sum_j (x_i + x_{i-1})\, \Delta x_i\, \Delta y_j + \frac{1}{2}\sum_i \sum_j (y_j + y_{j-1})\, \Delta x_i\, \Delta y_j$$

$$= \frac{1}{2}\sum_i (x_i^2 - x_{i-1}^2) \sum_j y_j + \frac{1}{2}\sum_j (y_j^2 - y_{j-1}^2) \sum_i \Delta_i$$

$$= \frac{1}{2}(b^2 - a^2)(d - c) + \frac{1}{2}(d^2 - c^2)(b - a)$$

$$= \frac{1}{2}(b - a)(d - c)(a + b + c + d)$$

Thus for any arbitrary partition P

$$L(P,f) \le \frac{1}{2}(b - a)(d - c)(a + b + c + d) \le U(P,f)$$

$\Rightarrow$ For all partitions P of R, a number $\frac{1}{2}(b-a)(d-c)(a+b+c+d)$, independent of P exists such that

$$\left[\frac{1}{2} - \varepsilon,\ \frac{1}{2} + \varepsilon';\ 0, 1\right],$$

$\Rightarrow$
$$\iint_R f\, dx\, dy = \frac{1}{2}(b - a)(d - c)(a + b + c + d)$$

CHAPTER 17

Example 5 (*b*). Evaluate $\iint (y-2x)\,dx\,dy$, over R = [1, 2; 3, 5].

■ Let $P_1 = \{1 = x_0, x_1, \ldots, x_n = 2\}$, and $P = \{3 = y_0, y_1, \ldots, y_m = 5\}$ be arbitrary partitions of [1, 2] and [3, 5] respectively. Then

$$P = P_1 \times P_2 = \{(x_i, y_j) : x_i \in P_1,\, y_j \in P_2\}$$

is an arbitrary partition of the rectangle R = [1, 2; 3, 5].

One each sub-rectangle, $\Delta R_{ij} = [x_{i-1}, x_i ; y_{j-1}, y_j]$, the numbers $(y_{j-1} - 2x_i)$ and $(y_j - 2x_{i-1})$ are the lower and the upper bounds of the function $y - 2x$.

Thus

$$L(P, f) = \sum_i \sum_j (y_{j-1} - 2x_i)\, \Delta x_i\, \Delta y_j$$

$$U(P, f) = \sum_i \sum_j (y_j - 2x_{i-1})\, \Delta x_i\, \Delta y_j$$

Also

$$y_{j-1} - 2x_i \le \frac{1}{2}(y_j + y_{j-1}) - (x_i + x_{i-1}) \le y_j - 2x_{i-1}$$

$\therefore$

$$L(P, f) \le \sum_i \sum_j \left\{\frac{1}{2}(y_j + y_{j-1}) + \frac{1}{2}(x_i + x_{i-1})\right\} \Delta x_i\, \Delta y_j \le U(P, f)$$

But;

$$\sum_i \sum_j \left\{\frac{1}{2}(y_j + y_{j-1}) - (x_i + x_{i-1})\right\} \Delta x_i\, \Delta y_j$$

$$= \frac{1}{2}\sum_i \sum_j (y_j + y_{j-1})\,\Delta y_j\, \Delta x_i - \sum_i \sum_j (x_i + x_{i-1})\, \Delta x_i\, \Delta y_j$$

$$= \frac{1}{2}\sum_j (y_j^2 - y_{j-1}^2) \sum_i \Delta x_i - \sum_i (x_i^2 - x_{i-1}^2) \sum_j \Delta_j$$

$$= \frac{1}{2}(5^2 - 3^2)(2-1) - (2^2 - 1^2)(5-3) = 2$$

Thus, for any arbitrary partition P of R,

$$L(P, f) \le 2 \le U(P, f).$$

so that for all partitions P of R, the number 2 (independent of P) exists which lies between $L(P, f)$ and $U(P, f)$.

Hence $$\iint_R (y - 2x)\, dx\, dy = 2$$

Ex. Show that the double integral of the following functions exists over R = [0, 1; 0, 1],

(*i*) $f(x, y) = x^2 + y^2$

(*ii*) $f(x, y) = \begin{cases} \sin(1/(x+y)), & (x, y) \ne (0, 0) \\ 0 & (x, y) = (0, 0) \end{cases}$

(*iii*) $f(x, y) = \begin{cases} 0, & \text{if } (x, y) = (1/m, 1/n), \text{ for some } m, n \in N \\ 1, & \text{otherwise} \end{cases}$

Example 6. A function f is defined on the rectangle $R = [0, 1; 0, 1]$ as follows:

$$f(x, y) = \begin{cases} \dfrac{1}{2}, & \text{when } y \text{ is rational} \\ x, & \text{when } y \text{ is irrational} \end{cases}$$

show that the double integral $\iint_R f(x, y)\, dx\, dy$ does not exist.

■ Let a partition

$$P_1 = [\{0 = x_0, x_1, \ldots, x_n = \frac{1}{2}, x_{n+1}, x_{n+2}, \ldots, x_{2n} = 1\}$$

divides the interval [0, 1] on x-axis into $2n$ equal sub-intervals, each of length $1/2n$.

Let a partition $P_2 = \{0 = y_0, y_1, y_2, \ldots, y_{2n} = 1\}$ divides [0, 1] on y-axis into $2n$ equal sub-intervals.

Lines drawn parallel to the axes through the points of the partitions P_1 and P_2 give rise to a partition P of R into $(2n)^2$ squares, ΔR_{ij}, each of area $1/4n^2$, so that

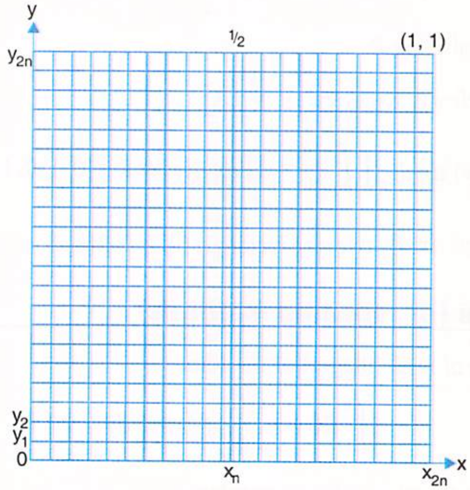


Fig. 4

$$x_i = \frac{i}{2n}, \Delta x_i = \frac{1}{2n}$$

$$\Delta R_{ij} = \Delta x_i \, \Delta y_j = \frac{1}{4n^2}$$

$$\sum_{j=1}^{2n} \Delta R_{ij} = \sum_{j=1}^{2n} \Delta x_i \, \Delta y_j \quad \Delta x_i = \frac{1}{2n}$$

$$\sum_{i=1}^{n} \sum_{j=1}^{2n} \Delta R_{ij} = \frac{1}{2}; \quad \sum_{i=n+1}^{2n} \sum_{j=1}^{2n} \Delta R_{ij} = \frac{1}{2}$$

$$\therefore \qquad L(P, f) = \sum_{i=1}^{2n}\sum_{j=1}^{2n} m_{ij}\,\Delta R_{ij} = \sum_{i=1}^{n}\sum_{j=1}^{2n} x_i\,\Delta R_{ij} + \sum_{i=n+1}^{2n}\sum_{j=1}^{2n}\frac{1}{2}\Delta R$$

$$= \sum_{i=1}^{n}\frac{i}{2n}\,\frac{1}{2n} + \frac{1}{2}\sum_{i=n+1}^{2n}\sum_{j=1}^{2n}\Delta R_{ij} = \frac{1}{4n^2}\,\frac{n(n+1)}{2} + \frac{1}{2}\cdot\frac{1}{2}$$

$$U(P,f) = \sum_{i=1}^{n}\sum_{j=1}^{2n}\tfrac{1}{2}\Delta R_{ij} + \sum_{i=n+1}^{2n}\sum_{j=1}^{2n} x_i\,\Delta R_{ij} = \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{2n}\Delta R_{ij} + \sum_{i=n+1}^{2n} x_i\cdot\frac{1}{2n}$$

$$= \frac{1}{2}\cdot\frac{1}{2} + \frac{1}{4n^2}\,\frac{n(3n+1)}{2}$$

$\therefore\ U(P, f) - L(P, f) = \dfrac{1}{4}$, for an arbitrary partition P,

i.e., for all partitions P of R,

$$U(P,f) - L(P,f) = \frac{1}{4} \nless \varepsilon\left(=\frac{1}{8}, \text{say}\right)$$

$\Rightarrow$ $\quad f$ is not integrable on R.

Ex. 1. Find $\iint_R (x + 2y)\,dx\,dy$, when $R = [1, 2; 3, 5]$.

Ex. 2. Evaluate $\iint_R (x - y)\,dx\,dy$, $\iint_R [x - y]\,dx\,dy$, for $R = [0, 1; 0, 1]$.

Ex. 3. Evaluate $\iint (x^2 + y)\,dx\,dy$, over the rectangle $[0, 1; 0, 2]$.

2.7 Iterated Integrals (*or repeated integrals*)

Definition. An *iterated integral* is an integral of the form

$$\int_a^b dx = \int_{\phi_1(x)}^{\phi_2(x)} f(x, y)\,dy$$

where ϕ_1 or ϕ_2 or both are functions of x or constants.

This means that for each fixed x between a and b, the integral

$$F(x) = \int_{\phi_1(x)}^{\phi_2(x)} f(x, y)\,dy$$

is evaluated, and then the integral $\int_a^b F(x)\,dx$.

$$\therefore \int_a^b F(x)\,dx = \int_a^b \left\{ \int_{\phi_1(x)}^{\phi_2(x)} f(x, y)\,dy \right\} dx$$

or

$$= \int_a^b dx \int_{\phi_1(x)}^{\phi_2(x)} f(x, y)\,dy \qquad \ldots(1)$$

The other repeated integral

$$\int_c^d dy \int_{\Psi_1(y)}^{\Psi_2(y)} f(x, y)\, dx \quad \text{or} \quad \int_c^d \left\{ \int_{\Psi_1(y)}^{\Psi_2(y)} f(x, y)\, dx \right\} dy \qquad ...(2)$$

is defined in the same way.

2.8 Reduction to Iterated (Repeated) Integrals (Calculation of a double integral over a rectangle)

Theorem 1. *Fubini's theorem. If a double integral,* $I = \iint_R f\, dx\, dy$ *exists over a rectangle* $R = [a, b;\, c, d]$, *and if* $\int_c^d f\, dy$ *also exists, for each fixed* x *in* $[a, b]$, *then the iterated integral* $\int_a^b dx \int_c^d f\, dy$ *exists and is equal to the double integral* I.

Let ε be any positive number.

Since the upper integral, I^u is the infimum of the upper sums, there exists a partition P of R such that

$$\sum_i \sum_j M_{ij}\, \Delta R_{ij} < I^u + \varepsilon$$

or

$$\sum_i \sum_j M_{ij}\, \Delta x_i\, \Delta y_j < I^u + \varepsilon \qquad ...(1)$$

Again, since for any fixed value of $x \in \Delta x_i$, M_{ij} is an upper bound (not necessarily the supremum) of f in Δy_j, therefore

$$\overline{\int_c^d} f\, dy \le \sum_j M_{ij}\, \Delta y_j, \text{ when } x \in \Delta x_i \qquad ...(2)$$

Again, from equation (2), $\sum_J M_{ij}\, \Delta y_j$ is an upper bound of the function

$$\phi(x) = \int_c^d f\, dy, \quad x \in \Delta x_i$$

therefore by the same reasoning as above, we get

$$\overline{\int_a^b} \phi(x)\, dx \le \sum_i \left(\sum_j M_{ij}\, \Delta y_j\right) \Delta x_i$$

or

$$\overline{\int_a^b} dx\, \overline{\int_c^d} f\, dy \le \sum_i \sum_j M_{ij}\, \Delta y_j\, \Delta x_i < I^u + \varepsilon \text{ [by (1)]} \qquad ...(3)$$

Also, since by hypothesis,

$$\underline{\int_c^d} f\, dy = \overline{\int_c^d} f\, dy = \int_c^d f\, dy$$

and ε is an arbitrary positive number, we get from equation (3),

$$\overline{\int_a^b} dx\, \overline{\int_c^d} f\, dy = \overline{\int_a^b} dx \int_c^d f\, dy \le I^u \qquad ...(4)$$

CHAPTER 17

By considering the lower integral I_l, we can similarly show that

$$\underline{\int}_a^b dx \overline{\int}_a^d f\ dy \geq I_l \qquad \ldots(5)$$

Again, since $I_l = I^u = I$ as the double integral exists, from (4) and (5), we get

$$I \leq \underline{\int}_a^b dx \int_c^d f\ dy \leq \overline{\int}_a^b dx \int_c^d f\ dy \leq I$$

$$\Rightarrow \qquad \underline{\int}_a^b dx \int_c^d f\ dy = \overline{\int}_a^b dx \int_c^d f\ dy = \iint_R f\ dx\ dy$$

Thus $\int_a^b \left[\int_c^d f\ dy\right] dx$ exists and equals $\iint_R f\ dx\ dy$.

Notes:

1. Similarly, if $\iint_R f\ dx\ dy$ and $\int_a^b f\ dx$ both exist, then the double integral can be expressed as

$$\iint_R f\ dx\ dy = \int_c^d dy \int_a^b f\ dx$$

2. Since ε is an arbitrary positive number, inequality (3) shows that

$$\overline{\int}_a^b dx \overline{\int}_c^d f\ dy \leq \overline{\iint}_R f\ dx\ dy$$

and similarly by considering the lower integrals,

$$\underline{\int}_a^b dx \underline{\int}_c^d f\ dy \geq \underline{\iint}_R f\ dx\ dy.$$

Remarks:

1. The theorem holds even if f has a finite number of discontinuities, or an infinite number of discontinuities lying on a finite number of lines $x = C_i$, $i = 1, 2, \ldots, m$ parallel to y-axis.

   For, the function $\phi(x) = \int_c^d f\ dy$ will be discontinuous at a finite number of points $C_1, C_2, \ldots, C_m$ only, and will therefore be still integrable.
   A similar remark holds for discontinuities parallel to x-axis.
2. If a double integral exists, then the two repeated integrals cannot exist without being equal. However, if the double integral does not exist, nothing can be said about the repeated integrals; they may or may not exist.
   Also one of the repeated integrals may exist or even that both may exist and be equal and yet the double integral may not exist, *i.e.*, the existence of one or of both of the repeated integrals is no guarantee for the existence of the double integral.
   However, if the two repeated integrals exist both are unequal, the double integral cannot exist.

ILLUSTRATIONS

1. A function is defined on a rectangle $[0, 1; 0, 1]$ as

$$f(x, y) = \begin{cases} \frac{1}{2} & \text{when } y \text{ is rational} \\ x, & \text{when } y \text{ is irrational} \end{cases}$$

The iterated integral $\int_0^1 dy \int_0^1 f\ dx = \frac{1}{2}$, but the other iterated integral does not exist.

For y rational

$$\int_0^1 f\ dx = \int_0^1 \frac{1}{2} dx = \frac{1}{2}$$

For y irrational

$$\int_0^1 f\ dx = \int_0^1 x\ dx = \frac{1}{2}$$

$$\therefore \qquad \int_0^1 dy \int_0^1 f\ dx = \int_0^1 \frac{1}{2} dx = \frac{1}{2}$$

Proceeding as for a function of a single variable, we may show that the integral $\int_0^1 f\ dy$ does not exist and so $\int_0^1 dx \int_0^1 f\ dy$ also does not exist.

It was shown earlier that the double integral $\iint_R f\ dx\ dy$ does not exist.

2. For the rectangle $R = [0, 1; 0, 1]$, let $f(x, y) = 1$ except when $x = \frac{1}{2}$, and let $f\left(\frac{1}{2}, y\right) = 1$, for irrational y, and $f\left(\frac{1}{2}, y\right) = -1$, for rational value of y.

The function has an infinite number of discontinuities along the line $x = \frac{1}{2}$.

From the rectangle R cut out the rectangle $\left[\frac{1}{2} - \varepsilon, \frac{1}{2} + \varepsilon'; 0, 1\right]$, where ε, ε' are positive and arbitrarily small, and let R_1 (consisting of two rectangles) be the remaining region.

Now

$$\iint_{R_1} f\ dx\ dy = \iint_{R_1} 1\ dx\ dy = 1 - \varepsilon - \varepsilon' \text{ (area of } R_1)$$

so that the double integral over R is (by definition) equal to unity.

Again

$$\iint_{R_1} f\ dx\ dy = \int_0^{\frac{1}{2}-\varepsilon} dx \int_0^1 1\ dy + \int_{\frac{1}{2}+\varepsilon'}^1 dx \int_0^1 1 \cdot dy = 1 - \varepsilon - \varepsilon'$$

$$\iint_{R_1} f\ dx\ dy = \int_0^1 dy \int_0^{\frac{1}{2}-\varepsilon} 1 \cdot dx + \int_0^1 dy \int_{\frac{1}{2}+\varepsilon'}^1 1 \cdot dx = 1 - \varepsilon - \varepsilon'$$

Proceeding to limits, when ε, ε' tend to zero, we get

$$\iint_R f\ dx\ dy = \int_0^1 dx \int_0^1 f\ dy = \int_0^1 dy \int_0^1 f\ dx = 1$$

It may also be noted that when $x=\frac{1}{2}$, the upper integral $\overline{\int_0^1} f\ dy = 1$ and the lower integral $\underline{\int_0^1} f\ dy = -1$, so that the integral $\int_0^1 f\ dy$ does not exist for $x=\frac{1}{2}$.

Thus the function $F(x) = \int_0^1 f(x, y)\ dy$ is discontinuous at $x=\frac{1}{2}$.

3. For the function f defined as :

$$f(x, y) = \begin{cases} 1/y^2, & \text{if } 0 < x < y < 1 \\ -1/x^2, & \text{if } 0 < y < x < 1 \\ 0, & \text{otherwise if } 0 \le x, y \le 1 \end{cases}$$

$$\int_0^1 dx \int_0^1 f\ dy \neq \int_0^1 dy \int_0^1 f\ dx$$

(0, 0) is a point of infinite discontinuity of f.

For $0 < x < 1$,

$$\int_0^1 f\ dy = -\int_0^x \frac{dy}{x^2} + \int_x^1 \frac{dy}{y^2} = -1$$

$$\therefore \qquad \int_0^1 dx = \int_0^1 f\ dy = \int_0^1 (-1)\ dx = -1$$

Again, for $0 < y < 1$,

$$\int_0^1 f\ dx = \int_0^y \frac{dx}{y^2} - \int_y^1 \frac{dx}{x^2} = 1$$

$$\therefore \qquad \int_0^1 dy \int_0^1 f\ dx = \int_0^1 1 \cdot dy = 1$$

Hence

$$\int_0^1 dx \int_0^1 f\ dy \neq \int_0^1 dy \int_0^1 f\ dx$$

As the two repeated integrals exist but are unequal, the double integral, $\iint_R f\ dx\ dy$ cannot exist.

Example 7. Show that

$$\int_0^1 dx \int_0^1 \frac{x^2 - y^2}{x^2 + y^2}\ dy = \int_0^1 dy \int_0^1 \frac{x^2 - y^2}{x^2 + y^2}\ dx$$

i.e., the change in the order of integration is permissible.

- Now

$$\int_0^1 \frac{x^2 - y^2}{x^2 + y^2}\, dy = \int_0^1 \left(\frac{2x^2}{x^2 + y^2} - 1\right) dy = 2x \tan^{-1}\frac{1}{x} - 1$$

$$\therefore \qquad \int_0^1 dx \int_0^1 \frac{x^2 - y^2}{x^2 + y^2}\, dy = \int_0^1 \left(2x \tan^{-1}\frac{1}{x} - 1\right) dx = 0$$

Again

$$\int_0^1 \frac{x^2 - y^2}{x^2 + y^2}\, dx = \int_0^1 \left(1 - \frac{2y^2}{x^2 + y^2}\right) dx = 1 - 2y \tan^{-1}\frac{1}{y}$$

$$\therefore \qquad \int_0^1 dy \int_0^1 \frac{x^2 - y^2}{x^2 + y^2}\, dx = \int_0^1 \left(1 - 2y \tan^{-1}\frac{1}{y}\right) dy = 0$$

Thus

$$\int_0^1 dx \int_0^1 \frac{x^2 - y^2}{x^2 + y^2}\, dy = 0 = \int_0^1 dy \int_0^1 \frac{x^2 - y^2}{x^2 + y^2}\, dx\,.$$

Example 8. Prove that

$$\int_0^1 dx \int_0^1 \frac{x - y}{(x + y)^3}\, dy = \frac{1}{2},\quad \int_0^1 dy \int_0^1 \frac{x - y}{(x + y)^3}\, dx = -\frac{1}{2}$$

Does the double integral $\iint_R \frac{x - y}{(x + y)^3}\, dx\, dy$ exist over $R = [0, 1; 0, 1]$?

- The integrand $(x - y)/(x + y)^3$ is bounded over the square $[0, 1; 0, 1]$ except at the origin which is a point of infinite discontinuity (The integral is actually an improper double integral).
  For $x \neq 0$,

$$\phi(x) \int_0^1 \frac{x - y}{(x + y)^3}\, dy = \int_0^1 \left[\frac{2x}{(x + y)^3} - \frac{1}{(x + y)^2}\right] dy = \frac{1}{(1 + x)^2}$$

and $\phi(0)$ does not exist, so that $\int_0^1 \phi(x)\, dx$ is an improper integral.

$$\therefore \qquad \int_0^1 dx \int_0^1 \frac{x - y}{(x + y)^3}\, dy = \int_0^1 \phi\, dx = \lim_{\lambda \to 0+} \int_\lambda^1 \frac{dx}{(1 + x)^2} = \lim_{\Lambda \to 0+}\left(\frac{1}{1+\lambda} - \frac{1}{2}\right) = \frac{1}{2}$$

Again, for $y \neq 0$,

$$\Psi(y) = \int_0^1 \frac{x - y}{(x + y)^3}\, dx = -\frac{1}{(x + y)^2}$$

and $\Psi(0)$ does not exist, so that proceeding as above

$$\int_0^1 dy \int_0^1 \frac{x-y}{(x+y)^3}\, dx = -\frac{1}{2}$$

Thus the two iterated integrals exist but are unequal and so the double integral cannot exist over R.

Note: The above example shows that change in the order of integration may not always give the same result. A sufficient condition under which such a change is permissible is that the double integral over the corresponding region exists. In the above case, the double integral fails to exist because the integrand is not bounded in a neighbourhood of the origin.

EXERCISE

1. Evaluate the following integrals:

   (*i*) $\iint_R (x^2 + 2y)\, dx\, dy, \quad R = [0, 1; 0, 2]$

   (*ii*) $\iint_R \frac{dx\, dy}{(x+y+1)^2}, \quad R = [0, 1; 0, 1]$

   (*iii*) $\iint_R \frac{x^2}{1+y^2}\, dx\, dy$ over the rectangle $[0, 1; 0, 1]$

   (*iv*) $\iint_R \frac{x-y}{x+y}\, dx\, dy, \quad R = [0, 1; 0, 1]$

   [In (*iv*) the integrand is bounded and $(0, 0)$ is the only point of discontinuity in R.]

2. Compute the double integrals of the following functions over R.

   (*i*) $x \sin(x+y), \quad R = [0, \pi; 0, \pi/2]$

   (*ii*) $x^2 y e^{xy}, \quad R = [0, 1; 0, 2]$

   (*iii*) $x^2 y \cos(xy^2), \quad R = [0, \pi/2; 0, 2]$

3. Let a function f be defined on $R = [0, 2; 0, 3]$ by

$$f(x, y) = \begin{cases} 3, & \text{if } x \text{ is rational} \\ y^2, & \text{if } x \text{ is irrational} \end{cases}$$

   Show that one of the iterated integrals exists, but the other does not.

4. Are the two iterated integrals

$$\int_1^\infty dx \int_1^\infty \frac{x-y}{(x+y)^3}\, dy, \text{ and } \int_1^\infty dy \int_1^\infty \frac{x-y}{(x+y)^3}\, dx \text{ equal?}$$

   If not, why so?

5. Show that $\phi(\lambda) = \int_\lambda^1 dx \int_\lambda^1 \frac{x-y}{(x+y)^3}\, dy$ is not continuous at $\lambda = 0$.

6. For $R = [a, b; c, d]$, show that

$$\iint_R \phi(x)\, \Psi(y)\, dx\, dy = \left[\int_a^b \phi(x)\, dx\right]\left[\int_c^d \Psi(y)\, dy\right]$$

7. For the function $f(x, y) = (y^2 - x^2)/(y^2 + x^2)^2$, show that

$$\int_\lambda^1 dx \int_\lambda^1 f\, dy = \int_\lambda^1 dy \int_\lambda^1 f\, dx = 0, \quad \text{if } 0 < \lambda < 1$$

but

$$\int_0^1 dx \int_0^1 f\, dy \neq \int_0^1 dy \int_0^1 f\, dx .$$

8. Prove the Dirichlet's formula :

$$\iint_R x^{p-1} y^{q-1}\, dx\, dy = \frac{\Gamma\left(\frac{p}{2}\right)\Gamma\left(\frac{q}{2}\right)}{\Gamma\left(\frac{p}{2}+\frac{q}{2}+1\right)},$$

whether (*i*) $p \geq 1, q \geq 1$, or (*ii*) $0 < p < 1, 0 < q < 1$, where R is the region bounded by the first quadrant of the circle $x^2 + y^2 = 1$.

ANSWERS

1. (*i*) 14/3 (*ii*) log 4/3 (*iii*) $\pi/12$ (*iv*) 0.
2. (*i*) $\pi - 2$ (*ii*) 2 (*iii*) $-\pi/16$.

2.9 Differentiation Under the Integral Sign

In this section, we shall prove that, under suitable conditions, '*the derivative of the integral and the integral of the derivative are equal*', and consequently, '*the two repeated integrals are equal for continuous functions.*'

Theorem 2. ***Leibnitz's rule.*** *If f is defined and continuous on the rectangle $R = [a, b; c, d]$, and if*

(*i*) *$f_x(x, y)$ exists and is continuous on the rectangle R, and*

(*ii*) $g(x) = \int_c^d f(x, y)\, dy$, for $x \in [a, b]$,

then g is differentiable on $[a, b]$, and

$$g'(x) = \int_c^d f_x(x, y)\, dy,$$

i.e.,

$$\frac{d}{dx}\left\{\int_c^d f(x, y)\, dy\right\} = \int_c^d \frac{\partial f(x,y)}{\partial x} dy$$

Since f_x exists on R, therefore, for each $y \in [c, d]$, and each $h \neq 0$, it follows by the Lagrange's mean-value theorem, that

$$f(x + h, y) - f(x, y) = hf_x(x + \theta h, y), \quad \text{for some } 0 < \theta < 1$$

Now f, being continuous, is integrable on $[c, d]$ for each $x \in [a, b]$, therefore $g(x)$ is a well-defined function on $[a, b]$.

$$\therefore \quad \frac{g(x+h)-g(x)}{h} = \frac{1}{h}\int_c^d \{f(x+h, y) - f(x, y)\}\, dy$$

$$= \int_c^d f_x(x+\theta h, y)\, dy, \; 0 < \theta < 1 \qquad ...(1)$$

Let $\varepsilon > 0$ be given. Then, by the continuity and so uniform continuity of f_x on R, $\exists\, \delta > 0$, such that if $(x, y), (x', y') \in R$ with $|x - x'| < \delta, |y - y'| < \delta$, then

$$|f_x(x, y) - f_x(x', y')| < \varepsilon/(d-c)$$

Let $0 < |h| < \delta$, then for each $y \in [c, d]$,

$$|f_x(x+\theta h, y) - f_x(x, y)| < \varepsilon/(d-c) \qquad ...(2)$$

From equations (1) and (2), we obtain

$$\left|\frac{g(x+h)-g(x)}{h} - \int_c^d f_x(x, y)\, dy\right| \le \int_c^d |f_x(x+\theta h, y) - f_x(x, y)|\, dy$$

$$< \frac{\varepsilon}{d-c}(d-c) = \varepsilon, 0 < |h| < \delta$$

Hence $g'(x) = \lim\limits_{h\to 0} \dfrac{g(x+h)-g(x)}{h} \displaystyle\int_c^d f_x(x, y)\, dy.$

Corollary 1. **(*General Leibnitz's rule*).** *If f satisfy the conditions of the above theorem, and if*

(*i*) $\phi, \Psi : [a, b] \to [c, d]$ *are both differentiable, and*

(*ii*) $g(x) = \displaystyle\int_{\phi(x)}^{\Psi(x)} f(x, y)\, dy$, *for* $x \in [a, b]$,

then g is differentiable on $[a, b]$, *and*

$$g'(x) = \int_{\phi(x)}^{\Psi(x)} f_x(x, y)\, dy + f(x, \Psi(x))\, \Psi'(x) - f(x, \phi(x))\, \phi'(x).$$

Let $G(x, \alpha, \beta) = \displaystyle\int_\alpha^\beta f(x, y)\, dy$, $x \in [a, b]$, and $\alpha, \beta \in \mathbf{R}$. Then, by the Leibnitz's rule,

$$\frac{\partial G}{\partial x} = \int_\alpha^\beta f_x(x, y)\, dy, \qquad ...(1)$$

which is continuous, by using continuity (uniform) of $f_x(x, y)$. Now, by the Fundamental theorem of calculus, we have

$$\frac{\partial G}{\partial \alpha} = -f(x, \alpha), \text{ and } \frac{\partial G}{\partial \beta} = f(x, \beta) \qquad ...(2)$$

which are continuous, by the continuity of $f(x, y)$.

By (*ii*), $g(x) = G(x, \phi(x), \Psi(x))$, for $x \in [a, b]$

Thus, G has continuous first order partial derivatives and is, therefore, differentiable. Moreover, the two functions ϕ and Ψ are differentiable on $[a, b]$. Therefore, applying the chain rule (§ 7.2 (I), Chapter 15) of differentiation of composite functions of g, we obtain

$$g'(x) = \frac{\partial G}{\partial x} + \frac{\partial G}{\partial \phi}\phi'(x) + \frac{\partial G}{\partial \Psi}\Psi'(x) \qquad ...(3)$$

From (1), (2) and (3), we get the required result.

Corollary 2. *If f is continuous on $R = [a, b; c, d]$, then*

$$\int_a^b \left\{\int_c^d f(x, y)\, dy\right\} dx = \int_c^d \left\{\int_a^b f(x, y)\, dx\right\} dy.$$

i.e., the two repeated (iterated) integrals are equal.

Define $g(x, z) = \int_c^z f(x, y)\, dy$, $x \in [a, b]$, then by the Fundamental theorem of calculus,

$$\frac{\partial g}{\partial z} = f(x, z), \text{ which is continuous, by the continuity of } f.$$

Let $G(z) = \int_a^b g(x, z)\, dx$, $z \in [c, d]$, then by the Leibnitz's rule

$$G'(z) = \int_a^b \frac{\partial g(x, z)}{\partial z}\, dx = \int_a^b f(x, z)\, dx,$$

which is again continuous, by using continuity (uniform) of f,

$$\therefore \qquad \int_c^d \left\{\int_a^b f(x, z)\, dx\right\} dz = \int_c^d G'(z)\, dz = G(d) - G(c)$$

[by the fundamental theorem of calculus]

$$= G(d)$$

$$= \int_a^b \left\{\int_c^d f(x, y)\, dy\right\} dx.$$

Hence the result follows.

Example 9. Show that

$$\int_0^{\pi/2} \log(1 - x^2 \sin^2\theta)\, d\theta = \pi \log(1 + \sqrt{1 - x^2}) - \pi \log 2, \text{ if } |x| < 1$$

■ The function $\log(1 - x^2 \sin^2\theta)$ is well-defined in the rectangle $[-1, 1; 0, \pi/2]$, and satisfies the conditions of the Leibnitz's rule.

Let $$g(x) = \int_0^{\pi/2} \log(1 - x^2 \sin^2\theta)\, d\theta, \quad |x| < 1 \qquad ...(1)$$

CHAPTER 17

By differentiating under the integral sign, w.r.t. x, we get

$$g'(x) = \int_0^{\pi/2} \frac{-2x \sin^2 \theta}{1 - x^2 \sin^2 \theta}\, d\theta = \frac{2}{x} \int_0^{\pi/2} \frac{(1 - x^2 \sin^2 \theta - 1)}{1 - x^2 \sin^2 \theta}\, d\theta, \quad x \neq 0$$

$$= \frac{\pi}{x} - \frac{2}{x} \int_0^{\pi/2} \frac{d\theta}{1 - x^2 \sin^2 \theta}, \text{ Put } \cot \theta = t$$

$$= \frac{\pi}{x} - \frac{2}{x} \int_0^{\infty} \frac{dt}{1 + t^2 - x^2},$$

$$= \frac{\pi}{x} - \frac{2}{x\sqrt{1 - x^2}} \tan^{-1} \frac{t}{\sqrt{1 - x^2}}\Bigg|_0^{\infty} = \frac{\pi}{x} - \frac{\pi}{x\sqrt{1 - x^2}}$$

Integrating w.r.t. x, we obtain

$$g(x) = \pi \log x - \pi \log \left\{\frac{1 - \sqrt{1 - x^2}}{x}\right\} + c, \text{ where } c \text{ is an arbitrary constant.}$$

$$= \pi \log \left\{\frac{x^2}{1 - \sqrt{1 - x^2}}\right\} + c$$

$$= \pi \log \left\{\frac{x^2(1 + \sqrt{1 - x^2})}{1 - (1 - x^2)}\right\} + c$$

$$= \pi \log (1 + \sqrt{1 - x^2}) + c$$

But $g(0) = 0$, by (1), therefore $c = -\pi \log 2$

Hence, $g(x) = \pi \log (1 + \sqrt{1 - x^2}) - \pi \log 2$, for $|x| < 1$.

Example 10. Prove that

$$\int_{\frac{1}{2}\pi - \alpha}^{\frac{1}{2}\pi} \sin \theta \cos^{-1} (\cos \alpha \operatorname{cosec} \theta)\, d\theta = \frac{\pi}{2} (1 - \cos \alpha)$$

■ Let $\quad g(\alpha) = \int_{\frac{1}{2}\pi - \alpha}^{\pi/2} \sin \theta \cos^{-1} (\cos \alpha \operatorname{cosec} \theta)\, d\theta \qquad \ldots(1)$

Applying the general Leibnitz's rule, we obtain

$$g'(\alpha) = \int_{\frac{1}{2}\pi - \alpha}^{\pi/2} \frac{\sin \alpha\, d\theta}{\sqrt{1 - \cos^2 \alpha \operatorname{cosec}^2 \theta}} + \sin\left(\frac{\pi}{2} - \alpha\right) \cos^{-1} \left(\cos \alpha \operatorname{cosec}\left(\frac{\pi}{2} - \alpha\right)\right)$$

$$= \int_{\frac{1}{2}\pi-\alpha}^{\pi/2} \frac{\sin\alpha \sin\theta \, d\theta}{\sqrt{\sin^2\theta - \cos^2\alpha}} + \cos\alpha \cos^{-1}(1) \qquad \text{...(2)}$$

$$= \int_0^{\sin\alpha} \frac{\sin\alpha \, dt}{\sqrt{\sin^2\alpha - t^2}} + 0, \text{ taking } \cos\alpha = t$$

$$= \sin\alpha \sin^{-1}\left(\frac{t}{\sin\alpha}\right)\Bigg|_0^{\sin\alpha} = \frac{\pi}{2}\sin\alpha$$

Integrating w.r.t. α, we get

$$g(\alpha) = -\frac{\pi}{2}\cos\alpha + c, \text{ where } c \text{ is an arbitrary constant.}$$

But $g(\pi/2) = \pi/2$, by (1), therefore $c = \pi/2$

Hence, $g(\alpha) = \pi/2\,(1 - \cos\alpha)$.

EXERCISE

1. If f is continuous on $R = [a, b; c, d]$, then prove that the function

$$g(x) = \int_c^d f(x, y)\, dy \text{ is continuous on } [a, b].$$

2. If f is continuous on $R = [a, b; c, d]$, then show that

$$F_{xy}(x, y) = f(x, y), \text{ for all } (x, y) \in R,$$

where
$$F(x, y) = \int_c^y dv \int_a^x f(u, v)\, du.$$

Deduce that
$$\int_a^b dx \int_c^d f\, dy = \int_c^d dy \int_a^b f\, dx.$$

3. If $|a| \le 1$, show that

$$\int_0^\pi \log(1 + a\cos x)dx = \pi \log\left[\frac{1}{2} + \frac{1}{2}\sqrt{1-a^2}\right]$$

hence or otherwise show that $\int_0^\pi \log(1 + \cos x)\, dx = -\pi \log 2$.

4. If $|x| < 1$, show that

(*i*) $\displaystyle\int_0^\pi \frac{\log(1 + x\cos y)}{\cos y}\, dy = \pi \sin^{-1} x$

(*ii*) $\displaystyle\int_0^{\pi/2} \log(1 - x^2\cos^2\theta)\, d\theta = \pi \log(1 + \sqrt{1 - x^2}) - \pi \log 2$.

5. Show that

(*i*) $\displaystyle\int_0^{\pi/2} \frac{\log(1 + x\sin^2 y)}{\sin^2 y}\, dy = \pi\left[\sqrt{(1 + x)} - 1\right],$

(*ii*) $\int_{-\pi/2}^{\pi/2} \log(1 + x \sin y)\, dy = \frac{\pi}{2}\log\left(1 + \sqrt{1 - x^2}\right), 0 \le x \le 1,$

(*iii*) $\int_0^{\pi/2} \frac{\log(1 + \cos\alpha \cos x)}{\cos x}\, dx = \frac{\pi^2 - 4\alpha^2}{8},$

(*iv*) $\int_0^{\pi/2} \log(a\cos^2\theta + b\sin^2\theta)\, d\theta = \pi \log\left[\tfrac{1}{2}\left(\sqrt{a} + \sqrt{b}\right)\right]; a, b > 0,$

(*v*) $\int_0^{\pi/2} \log\left(\frac{a + b\sin\theta}{a - b\sin\theta}\right) \operatorname{cosec}\theta\, d\theta = \pi \sin^{-1}\frac{b}{a}, a > b.$

6. If $g(x) = \int_{\sin x}^{e^x} \sqrt{1 + y^3}\, dy$, find $g'(x)$.

7. Show that

(*i*) $\int_0^{x^2} \tan^{-1}(y/x^2)\, dy = (\pi - 2\log 2)\, x^2/4,$

(*ii*) $\int_0^{a} \frac{\log(1 + ax)}{1 + x^2}\, dx = \frac{1}{2}\log(1 + a^2)\tan^{-1} a, a > 0.$

8. Starting from a suitable integral, show that

(*i*) $\int_0^{x} \frac{dx}{(x^2 + a^2)^2} = \frac{1}{2a^3}\tan^{-1}\left(\frac{x}{a}\right) + \frac{x}{2a^2(x^2 + a^2)},$

(*ii*) $\int_0^{\pi/2} \frac{\sin^2 x\, dx}{(a^2\sin^2 x + b^2\cos^2 x)^2} = \frac{\pi}{4a^3 b}.$

3. DOUBLE INTEGRALS OVER A REGION

So far we have studied the double integrals over rectangles. In this section, we shall study double integration over a closed bounded domain (region E).

The procedure and the results will be similar to those for integrals over rectangles except for some verbal changes.

3.1 The Area of a Closed Bounded Domain E

Since E is bounded, a rectangle R exists which completely encloses E. Let P be a partition of R.

Let $L(P)$ denote the sum of the areas of the sub-rectangles which consist entirely of points of E, and $U(P)$ denote the sum of the areas of the sub-rectangles which have at least one point in common with E.

Clearly, $L(P) \le U(P)$

To every partition P of R, there corresponds a pair of sums $L(P)$ and $U(P)$. Clearly the sets of these sums are bounded.

The supremum of the set of sums $L(P)$ is called the *inner area of E* and the infimum of the set of sums $U(P)$ is called the *outer area* of E.

The region E is said to possess an area if the inner and the outer area of E are equal and the common value is called the area of E, generally denoted by the symbol S.

3.2 Double Integral over a Region E

Let a function of two variables be single-valued and bounded on a region E of area S.

Let a partition P, consisting of a finite number of curves, divide the region E into n sub-regions of elementary areas $\Delta S_1, \Delta S_2, \ldots, \Delta S_n$. These sub-regions may be formed by drawing curves that cover the area like a net. Clearly the sum of these elementary areas is $\sum_{i=1}^{n} \Delta S_i = S$.

Let M, m and M_i, m_i denote the upper and the lower bounds of f in E and ΔS_i respectively. From the sums

$$U(P, f) = \sum_i M_i \, \Delta S_i, \text{ and } L(P, f) = \sum_i m_i \, \Delta S_i$$

called the *upper* and the *lower sums* respectively.

As before

$$mS \le L(P, f) \le U(P, f) \le MS \qquad \ldots(1)$$

so that the two sets of sums are bounded and each has the infimum and the supremum. The infimum of the set of upper sums is called the *upper integral*, I^u, and the supremum of the set of lower sums is called the *lower integral*, I_l, of f over E, denoted as

$$\left.\begin{aligned} I^u &= \overline{\int}_E f \, dS \quad \text{or} \quad \overline{\iint}_E f \, dx \, dy \\ I_l &= \underline{\int}_E f \, dS \quad \text{or} \quad \underline{\iint}_E f \, dx \, dy \end{aligned}\right\} \qquad \ldots(2)$$

When the two integrals are equal, f is said to be *integrable* and the common value I is called the *double integral* of f over E, denoted as

$$I = \iint_E f \, dx \, dy \int_E f \, dS \qquad \ldots(3)$$

Note: Taking $f = 1$, we at once deduce that the area S of the plane region E in the xy-plane is given by

$$S = \iint_E dx \, dy$$

$dx \, dy$ being an elementary area of the region, which when expressed in polar coordinates becomes $r \, dr \, d\theta$.

Integral as a limit of sums. Let (ξ_i, η_i) be any point of ΔS_i. The norm $\mu(P)$ is the maximum diameter of the sub-regions produced by a partition P.

The limit $\lim_{\mu(P)\to 0} S(P, f) = \lim_{\mu(P)\to 0} \sum f(\xi, \eta_i) \, \Delta S_i$

if it exists for all positions of (ξ_i, η_i) in ΔS_i, and all partitions P of E are called the double integral of f over E.

$$\therefore \qquad \lim_{\mu(P)\to 0} S(P, f) = \iint_E f \, dS = \iint_E f \, dx \, dy \qquad \ldots(4)$$

CHAPTER 17

For simple regions, since the increase in the number of sub-regions reduces their size, $\mu(P) \to 0$ as $n \to \infty$. Therefore we may state the above result as

$$\lim_{n\to\infty} S(P, f) = \iint_E f\,dx\,dy \qquad \text{...(5)}$$

Note: It may be seen that $f(\xi_i, \eta_i)\,\Delta S_i$ represents the volume of a solid cylinder with base ΔS_i on the xy-plane, bounded at the top by the surface $z = f(x, y)$ and having generators (sides) parallel to the axis of z. Thus in the limit $(n \to \infty)$,

$$\lim_{\mu(P)\to 0} S(P, f) = \iint_E f\,dx\,dy$$

represents the volume of a cylindrical solid bounded above by the surface $z = f(x, y)$, below by the plane base E (projection of $z = f(x, y)$) in the xy-plane and having sides parallel to the z-axis.

By proceeding on the same lines, the results for the double integrals over rectangles (§ 2 to 2.1) can be easily established for double integrals over plane regions. We shall give formal proof of one of the theorems equivalent to that of § 2.6.

Theorem 3. *A bounded function f on a region E, having an infinite number of discontinuities lying on a finite number of smooth curves, is integrable on E.*

We first prove the theorem for the case in which the given function f is continuous at all points on a smooth curve, say C.

Let ε be any positive number.

Draw two curves C_1 and C_2 which cut out the line of discontinuity from the region E, so that the function is continuous in the remaining parts E_1 and E_2, of the region E.

The curves C_1 and C_2 may be drawn so close to C that the area they cut out from E will be as small as we please, say less than $\varepsilon/4M$ where M is the upper bound of f in E. Since the oscillation of f is less than or equal to $2M$, the contribution to the oscillatory sum from this area is therefore less than $2M(\varepsilon/4M)$, *i.e.*, less than $\frac{1}{2}\varepsilon$.

In the regions E_1 and E_2, f is continuous and therefore a partition P exists such that the contribution to the oscillatory sum $U(P, f) - L(P, f)$ from these regions is less than $\frac{1}{2}\varepsilon$. Thus a partition P of the area exists for which $U(P, f) - L(P, f) < \varepsilon$, and thus f is integrable over E.

The other lines of discontinuity may be dealt with similarly.

Remark: If a function is integrable over a region E (even if a rectangle), the values of f may be arbitrarily changed at isolated points in E, or at all points on a finite number of curves without affecting the existence or the value of the integral, provided the values of f are finite. It would be sufficient to consider these isolated points or the curves, among the discontinuities of f.

The truth of the statement is evident from the nature of the above proof.

Integrability over a region (bounded domain) can also be locked upon as follows:

3.3 Integrability over a Bounded Domain

When the given region E is bounded, a rectangle R can be found which completely enclosed E. Let us define a new function F on R.

$$F(x, y) = \begin{cases} f(x, y), & \text{at points of } E \\ 0, & \text{elsewhere} \end{cases}$$

A function f is said to be integrable on E, if F is integrable on R, and then

$$\iint_E f\, dx\, dy = \iint_R F\, dx\, dy.$$

Remark: If a function f is continuous on a domain E, bounded by a finite number of continuous curves of the form $y = \phi(x)$, $x = \Psi(y)$, etc., then $\iint_E f\, dx\, dy$ exists.

The result follows from the fact that the only possible points of discontinuity of F defined as above, are the points of the curves $y = \phi(x)$, etc.

3.4 Reduction to Iterated Integrals (*Calculation of a double integral over a closed domain*)

If a double integral $\iint_E f\, dx\, dy$ *exists for a function f defined on a closed regular domain E bounded by the curves*

$$y = \phi(x),\ y = \Psi(x);\ \ x = a, x = b,$$

where ϕ, Ψ *are continuous, and* $\phi(x) \le \Psi(x)$, *for all* $x \in [a, b]$, *and if the integral* $\int_{\phi(x)}^{\Psi(x)} f\, dy$ *exists for each fixed point x in* $[a, b]$, *then the iterated integral* $\int_a^b dx \int_{\phi(x)}^{\Psi(x)} f\, dy$ *also exists, and*

$$\iint_E f\, dx\, dy = \int_a^b dx \int_{\phi(x)}^{\Psi(x)} f\, dy \qquad \text{...(1)}$$

Let a rectangle $R = [a, b; c, d]$ enclose E.

Let us define a function F over R such that

$$F(x, y) = \begin{cases} f(x, y), & \text{at points of } E \\ 0, & \text{outside } E \end{cases}$$

Then

$$\iint_E f\, dx\, dy = \iint_R F\, dx\, dy = \int_a^b dx \int_c^d F\, dy \quad \text{(by Fubini's theorem)}$$

$$= \int_a^b dx \left\{ \int_c^{\phi(x)} F\, dy + \int_{\phi(x)}^{\Psi(x)} F\, dy + \int_{\Psi(x)}^{d} F\, dy \right\}$$

$$= \int_a^b dx \int_{\phi(x)}^{\Psi(x)} F\, dy = \int_a^b dx \int_{\phi(x)}^{\Psi(x)} f\, dy$$

The other two integrals vanish in view of the definition of F.

CHAPTER 17

Corollary. If instead, the domain is bounded by continuous curves

$$x = \phi_1(y),\ x = \Psi_1(y);\ y = c,\ y = d$$

where

$$\phi_1(y) \le \Psi_1(y),\ \forall\ y \in [c, d]$$

and $\iint_E f\ dx\ dy$ on E and $\int_{\phi_1(y)}^{\Psi_1(y)} f\ dx$ exists for each fixed y in $[c, d]$, then the iterated integral $\int_c^d dy \int_{\phi_1(y)}^{\Psi_1(y)} f\ dx$ exists, and

$$\iint_E f\ dx\ dy = \int_c^d dy \int_{\phi_1(y)}^{\Psi_1(y)} f\ dx \qquad ...(2)$$

Notes:

1. If a function f is continuous, then the double integral $\iint_E f\ dx\ dy$ is expressed in terms of iterated integrals by (1) or (2) depending on whether the domain is quadratic with respect to y-axis or x-axis.
2. The domain E in the above theorem is taken to be quadratic (regular) with respect to one of the coordinate axes. If E is not quadratic with respect to y-axis, result (1) cannot be used to express a double integral in terms of iterated integrals. However, if it is quadratic with respect to x-axis, result (2) can be used, and vice versa.

   If the domain E is not quadratic with respect to either axes, see if we can manage to partition E into a finite number of quadratic domains. Then, by evaluating the double integral over each of these sub-domains by means of iterated integrals, and adding the results, we get the required integral over E.

Example 11. Evaluate $\iint y\ dx\ dy$ over the part of the plane bounded by the lines $y = x$ and the parabola $y = 4x - x^2$.

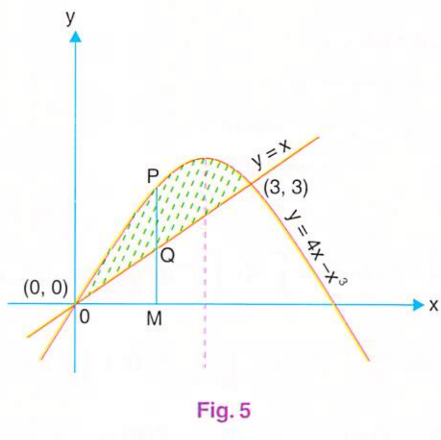


Fig. 5

- The line and the parabola intersect at the points (0, 0) and (3, 3), the domain of integration being the shaded region in the figure. Any line parallel to y-axis cuts the boundary of the region into two points, such as P, Q. Thus

$$\iint_E y\,dx\,dy = \int_0^3 dx \int_x^{4x-x^2} y\,dy = \tfrac{1}{2}\int_0^3 (x^4 - 8x^3 + 15x^2)\,dx = \frac{54}{5}$$

Example 12. Find the value of $\iint_E e^{y/x}\, dS$ if the domain E of integration is the triangle bounded by the straight lines $y = x$, $y = 0$ and $x = 1$.

- To avoid integration of $e^{y/x}$ with respect to x, we use

$$\iint_E e^{y/x}\, dS = \int_0^1 dx \int_0^x e^{y/x}\, dy = \int_0^1 x(e-1)\, dx = \frac{e-1}{2}$$

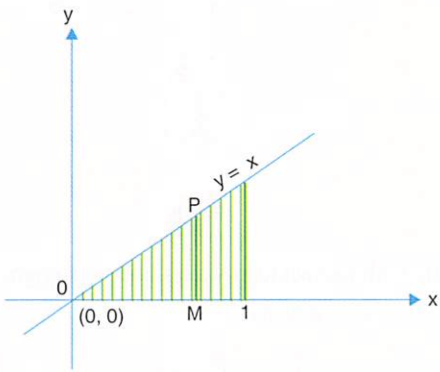


Fig. 6

Example 13. Change the order of integration in

$$I = \int_0^1 dx \int_x^{\sqrt{x}} f(x, y)\, dy.$$

- The line $y = x$ and the parabola $y = \sqrt{x}$ cut at (0, 0) and (1, 1). The domain of integration is the shaded region. Any line parallel to x-axis cuts the region into two points such as Q and P.

$$\therefore \qquad I = \int_0^1 dy \int_{y^2}^{y} f(x, y)\, dx$$

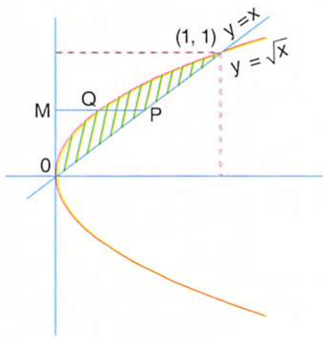


Fig. 7

Example 14. Evaluate the double integral $\int_E e^{x+y}\, dx\, dy$, when E is the domain which lies between two squares of sides 2 and 4, with centre at the origin and sides parallel to the axes.

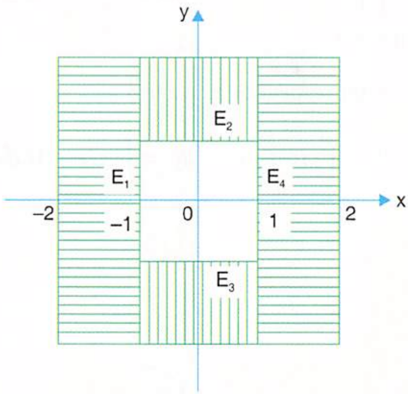


Fig. 8

- Domain E is not quadratic with respect to any axes but the straight lines $x = -1$, $x = 1$ divide it into four quadratic domains E_1, E_2, E_3 and E_4.

$$\therefore \qquad \iint_E e^{x+y}\, dx\, dy = \int_{-2}^{-1} dx \int_{-2}^{2} e^{x+y}\, dy + \int_{-1}^{1} dx \int_{1}^{2} e^{x+y}\, dy$$

$$+ \int_{-1}^{1} dx \int_{-2}^{-1} e^{x+y}\, dy + \int_{1}^{2} dx \int_{-2}^{2} e^{x+y}\, dy$$

$$= \int_{-2}^{-1} (e^{x+2} - e^{x-2})\, dx + \int_{-1}^{1} (e^{x+2} - e^{x+1})\, dx$$

$$+ \int_{-1}^{1} (e^{x-1} - e^{x-2})\, dx + \int_{1}^{2} (e^{x+2} - e^{x-2})\, dx$$

$$= 4 \sin h\, 3 \sin h\, 1.$$

Example 15. Evaluate

$$\iint_R f(x, y)\, dx\, dy,$$

over the rectangle

$$R = [0, 1; 0, 1],$$

where

$$f(x, y) = \begin{cases} x + y, & \text{if } x^2 < y < 2x^2 \\ 0, & \text{elsewhere} \end{cases}$$

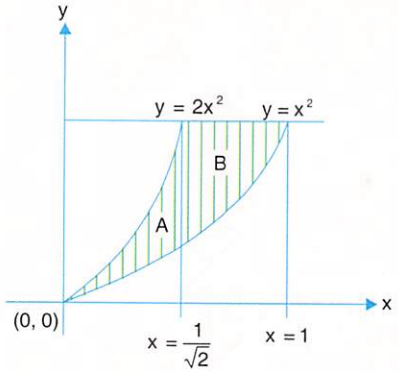


Fig. 9

■ For non-zero values of the function $f(x, y)$, the domain R is divided into two domains A and B, we obtain

$$\iint_R f(x, y)\,dx\,dy = \iint_A (x+y)\,dx\,dy + \iint_B (x+y)\,dx\,dy$$

$$= \int_0^{1/\sqrt{2}} dx \int_{x^2}^{2x^2} (x+y)\,dy + \int_{1/\sqrt{2}}^{1} dx \int_{x^2}^{1} (x+y)\,dy$$

$$= \int_0^{1/\sqrt{2}} \left(x^3 + \frac{3}{2}x^4\right) dx + \int_{1/\sqrt{2}}^{1} \left(x + \frac{1}{2} - x^3 - \frac{x^4}{2}\right) dx$$

$$= (21 - 8\sqrt{2})/40.$$

Example 16. Evaluate $\iint_R [x+y]\,dx\,dy$, over the rectangle $R = [0, 1; 0, 2]$, where $[x+y]$ denotes the greatest integer less than or equal to $(x+y)$.

■ We have, for $(x, y) \in R$.

$$[x+y] = \begin{cases} 0, & \text{if } 0 \le x+y < 1 \\ 1, & \text{if } 1 \le x+y < 2 \\ 2, & \text{if } 2 \le x+y < 3 \end{cases}$$

The domain of integration R is divided into three domains, therefore we have

$$\iint_R [x+y]\,dx\,dy = \int_0^1 dx \int_0^{1-x} [x+y]\,dy + \int_0^1 dx \int_{1-x}^{2-x} [x+y]\,dy + \int_0^1 dx \int_{2-x}^{2} [x+y]\,dy$$

$$= \int_0^1 dx \int_0^{1-x} 0\,dy + \int_0^1 dx \int_{1-x}^{2-x} 1\,dy + \int_0^1 dx \int_{2-x}^{2} 2\,dy = 0 + 1 + 1 = 2.$$

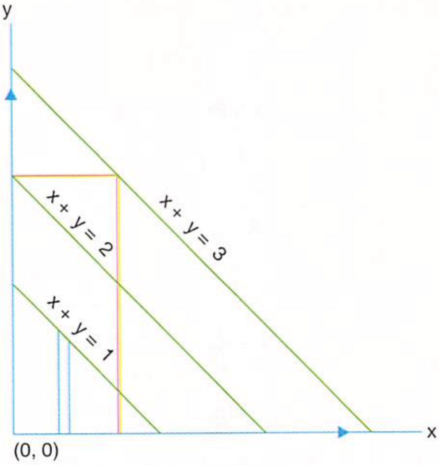


Fig. 10

Example 17. Prove that

$$\iint_R \sqrt{|y - x^2|}\, dx\, dy = (3\pi + 8)/6,$$

where $R = [-1, 1; 0, 2]$.

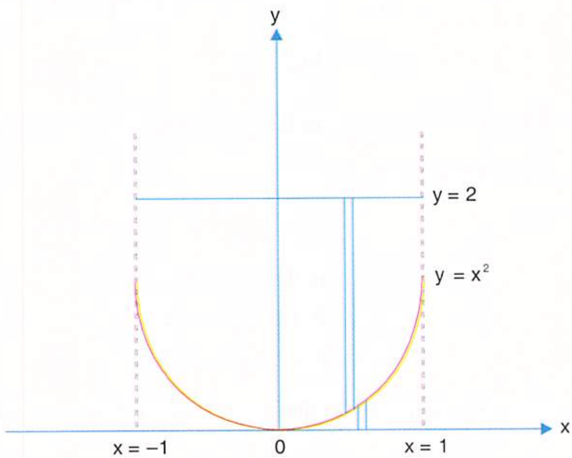


Fig. 11

■ The domain R is divided into two parts, as shown in the Figure, therefore we obtain

$$\iint_R \sqrt{|y - x^2|}\, dx\, dy = \int_{-1}^{1} dx \int_0^{x^2} \sqrt{(x^2 - y)}\, dy + \int_{-1}^{1} dx \int_{x^2}^{2} \sqrt{(y - x^2)}\, dy$$

$$= \frac{2}{3}\int_{-1}^{1} x^3\, dx + \frac{2}{3}\int_{-1}^{1} (2 - x^2)^{3/2}\, dx$$

$$= 0 + \frac{4}{3}\int_{0}^{1} (2 - x^2)^{3/2}\, dx$$

$$= \frac{16}{3}\int_{0}^{\pi/4} \cos^4\theta\, d\theta, \text{ taking } x = \sqrt{2}\sin\theta$$

$$= \frac{4}{3}\int_{0}^{\pi/4} (1 + \cos 2\theta)^2\, d\theta = \frac{4}{3} + \frac{\pi}{2}.$$

EXERCISE

(The reader is advised to draw the figure in each case)

1. Change the order of integration:

(*i*) $\int_{-1}^{1} dx \int_{0}^{\sqrt{(1-x^2)}} f\, dy$

(*ii*) $\int_{0}^{a} \left\{ \int_{x}^{\sqrt{(2^a x - x^2)}} f\, dy \right\} dx$

(*iii*) $\int_{-1/\sqrt{2}}^{1/\sqrt{2}} \left\{ \int_{-\sqrt{(1-2x^2)}}^{\sqrt{(1-2x^2)}} f\, dy \right\} dx$

(*iv*) $\int_{1}^{2} \left\{ \int_{x}^{2x} f\, dy \right\} dx$

(*v*) $\int_{0}^{1} dy \int_{\sqrt{y}}^{2-y} f\, dx$

(*vi*) $\int_{0}^{2a} dx \int_{\sqrt{(2ax-x^2)}}^{\sqrt{2ax}} f\, dy.$

2. By changing the order of integration, show that

(*i*) $\int_{0}^{1} dx \int_{0}^{x} f\, dy + \int_{1}^{2} dx \int_{0}^{2-x} f\, dy = \int_{0}^{1} dy \int_{y}^{2-y} f\, dx.$

(*ii*) $\int_{0}^{1} \left\{ \int_{0}^{x^2} f\, dy \right\} dx + \int_{1}^{3} \left\{ \int_{0}^{(3-x)/2} f\, dy \right\} dx = \int_{0}^{1} \left\{ \int_{\sqrt{y}}^{3-2y} f\, dx \right\} dy.$

(*iii*) $\int_{a^2/b}^{a} dy \int_{a^2/y}^{b} f\, dx + \int_{a}^{b} dy \int_{y}^{b} f\, dx = \int_{a}^{b} dx \int_{a^2/x}^{x} f\, dy.$

(*iv*) $\int_{0}^{1} dx \int_{0}^{x^{2/3}} f\, dy + \int_{1}^{2} dx \int_{0}^{1-\sqrt{(4x-x^2-3)}} f\, dy = \int_{0}^{1} dy \int_{y^{3/2}}^{2-\sqrt{(2y-y^2)}} f\, dx.$

(*v*) $\int_{0}^{(a/c)\sqrt{c^2-a^2}} dx \int_{0}^{ab/c} f\, dy + \int_{(a/c)\sqrt{c^2-a^2}}^{a} dx \int_{0}^{(b/a)\sqrt{a^2-x^2}} f\, dy = \int_{0}^{ab/c} dy \int_{0}^{(a/b)\sqrt{b^2-y^2}} f\, dx.$

3. Evaluate

(*i*) $\iint x^3 y^2\, dx\, dy$, over the circle $x^2 + y^2 \le a^2$.

(*ii*) $\iint (x^2 + y^2)\, dx\, dy$, over the domain bounded by $y = x^2$ and $y^2 = x$.

(*iii*) $\iint_E \frac{x^2}{y^2}\, dx\, dy$, E is bounded by $x = 2$, $y = x$, $xy = 1$.

(iv) $\iint \cos(x+y)\, dx\, dy$, over the domain enclosed by $x=0,\ y=\pi,\ y=x$.

(v) $\iint_E xe^{x^2-y^2}\, dx\, dy$, where E is the closed region bounded by the lines $y=x$, $y=x-1$, $y=0$, and $y=1$.

(vi) $\iint (x^2+y^2)\, dx\, dy$, over the region bounded by $xy=1$, $y=0$, $y=x$, and $x=2$.

4. By changing the order of integration, prove that

(i) $$\int_0^1 dx \int_x^{1/x} \frac{y^2\, dy}{(x+y)^2\sqrt{1+y^2}} = \left(2\sqrt{2}-1\right)/2$$

(ii) $$\int_0^1 dx \int_x^{1/x} \frac{y\, dy}{(1+xy)^2(1+y^2)} = (\pi-1)/4$$

(iii) $$\int_0^1 dx \int_0^{\sqrt{(1-x^2)}} \frac{dy}{(1+e^y)\sqrt{(1-x^2-y^2)}} = (\pi/2)\log[2e/(1+e)].$$

(iv) $$\int_0^{2a} dx \int_0^{\sqrt{(2ax-x^2)}} \frac{x(x^2+y^2)\,\phi'(y)\, dy}{\sqrt{4a^2x^2-(x^2+y^2)^2}} = \pi a^2[\phi(a)-\phi(0)]$$

5. Show that

$$\int_0^a dx \int_0^x f(x,y)\, dy = \int_0^a dy \int_y^a f(x,y)\, dx$$

and hence deduce Dirichlet's formula

$$\int_0^t dx \int_0^x \phi(y)\, dy = \int_0^t (t-y)\,\phi(y)\, dy.$$

6. Show that

$$\iint_E x^2y^2\sqrt{1-x^3-y^3}\, dx\, dy = \frac{4}{135}$$

Domain E is bounded by $x^3+y^3=1$, and $x \geq 0,\ y \geq 0$.

7. Evaluate by changing the order of integration

$$\int_2^4 \left\{\int_{4/x}^{(20-4x)/(8-x)} (4-y)\, dy\right\} dx.$$

8. Set the limits of integration in the double integral $\iint_E f(x,y)\, dx\, dy$, where E is the annulus bounded by the circles $x^2+y^2=1$, $x^2+y^2=4$.

[*Hint*. Divide E into four sub-domains by drawing the vertical tangents of the inner circle.]

ANSWERS

1. (i) $\int_0^1 dy \int_{-\sqrt{(1-y^2)}}^{\sqrt{(1-y^2)}} f\, dx$ (ii) $\int_0^a dy \int_{a-\sqrt{(a^2-y^2)}}^{y} f\, dx$

(iii) $\int_{-1}^1 dy \int_{-\sqrt{\{(1-y^2)/2\}}}^{\sqrt{\{(1-y^2)/2\}}} f\, dx$ (iv) $\int_1^2 dy \int_1^y f\, dx + \int_2^4 dy \int_{y/2}^2 f\, dx$

(v) $\int_0^1 dx \int_0^{x^2} f\, dy + \int_1^2 dx \int_0^{2-x} f\, dy$

(vi) $\int_0^a dy \int_{y^2/2a}^{a-\sqrt{(a^2-y^2)}} f\, dx + \int_0^a dy \int_{a+\sqrt{(a^2-y^2)}}^{2a} f\, dx + \int_a^{2a} dy \int_{y^2/2a}^{2a} f\, dx$

3. (i) 0 (ii) 6/35 (iii) 9/4 (iv) –2 (v) $(e^3 - e - 2)/4$ (vi) 47/24

7. $12 - 16 \log 2$.

4. GREEN'S THEOREM

We shall now discuss Green's Theorem, which provides a formula *connecting a line integral along a closed contour with a double integral over the domain bounded by that contour*. The theorem is due to an English mathematician, G. Green (1793 – 1841). In this section, we shall give its version in plane, that in a space of three dimensions will be discussed later.

Green's Theorem (*in R^2*). *If a domain E, regular with respect to both the axes, is bounded by a contour C, and f and g are two single-valued functions which along with their partial derivatives f_y, and g_x are continuous on E, then*

$$\iint_E \left(\frac{\partial g}{\partial x} - \frac{\partial f}{\partial y}\right) dx\, dy = \int_C (f\, dx + g\, dy)$$

where the line integral is taken in the positive direction.

Let us first consider a function f which, alongwith its partial derivative f_y, is continuous on a region E, regular with respect to y-axis. Let E be bounded by contour C, consisting of the curves $y = \phi(x),\ y = \Psi(x),\ x = a,\ x = b,$ such that

$$\phi(x) \le \Psi(x), \quad \forall\ x \in [a, b]$$

we have

$$\iint_E f_y(x, y)\, dx\, dy = \int_a^b dx \int_{\phi(x)}^{\Psi(x)} f_y(x, y)\, dy$$

$$= \int_a^b f(x, \Psi(x))\, dx - \int_a^b f(x, \phi(x))\, dx.$$

The two integrals on the right are line integrals along the contour from A to B and from C to D, respectively. (The portions BC and DA of the contour coincide with the lines $x = a$ and $x = b$ respectively.) Also contour

$$C = C_1 + C_2 + C_3 + C_4$$

Now

$$\int_a^b f(x, \Psi(x))\, dx = \int_{-C_1} f(x, y)\, dx = -\int_{C_1} f(x, y)\, dx$$

$$\int_{C_2} f(x, y)\, dx = 0 = \int_{C_4} f(x, y)\, dx\,.$$

$$\int_a^b f(x, \phi(x))\, dx = \int_{C_3} f(x, y)\, dx$$

$$\therefore \qquad \iint_E f_y\, dx\, dy = -\int_{C_1} f\, dx - \int_{C_2} f\, dx - \int_{C_3} f\, dx - \int_{C_4} f\, dx$$

$$= -\int_C f\, dx \qquad \ldots(1)$$

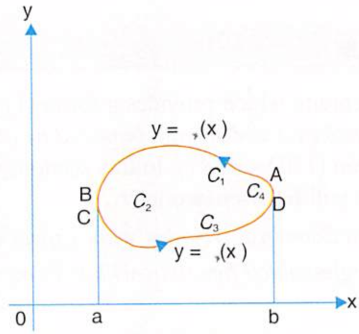


Fig. 12

In order to extend the result to any region, let us first suppose the region E to be piecewise regular with regard to the y-axis, *i.e.*, E can be split up into a finite number of sub-regions, $E_1, E_2, \ldots\ldots, E_n$, the contour of each of which is cut in at most two points by a line parallel to y-axis.

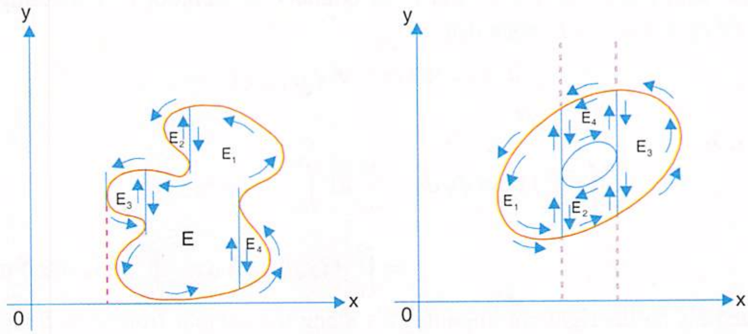


Fig. 13

Since result (1) holds for each sub-region, let us apply it to all the sub-regions and add. The double integrals add up to the double integral on the whole region E. Since the line integrals along the partition lines cancel each other because along each line the integral is taken twice in opposite directions, therefore the line integrals (in positive sense) along the contours of sub-regions also add up to the line integral along the contour C of the whole region E.

Hence for any region E, piecewise regular with respect to y-axis, we have

$$\iint_E f_y\, dx\, dy = -\int_C f\, dx \qquad \ldots(2)$$

It can similarly be proved that for any region E, piecewise regular with respect to x-axis,

$$\iint_E g_x \, dx \, dy = \int_C g \, dy \qquad ...(3)$$

Thus for any region E, which is regular with respect to both the axes, we have

$$\iint_E \left(\frac{\partial g}{\partial x} - \frac{\partial f}{\partial y} \right) dx \, dy = \int_C f \, dx + g \, dy \qquad ...(4)$$

Hence the theorem.

Note: The theorem holds even when the domain is enclosed by only two curves, $y = \phi(x)$ and $y = \Psi(x)$, between $x = a$ and $x = b$, the contours C_2 and C_4 reducing to zero.

Corollary. If $f(x, y) = y$, we see from (2) that the area of a domain which is piecewise regular with respect to y-axis

$$= -\int_C y \, dx. \qquad ...(5)$$

Similarly putting $g(x, y) = x$, we see from (3) that the area of a domain, piecewise regular with respect to x-axis

$$= \int_C x \, dy. \qquad ...(6)$$

By adding the above results, or putting $g(x, y) = x$ and $f(x, y) = -y$, we see from (4) that

$$= \frac{1}{2} \int_C (x \, dy - y \, dx) = \iint_E dx \, dy = \text{area of } E$$

Thus the area of a domain E (with contour C), regular with respect to both the axes

$$= \frac{1}{2} \int_C (x \, dy - y \, dx). \qquad ...(7)$$

ILLUSTRATION

Use line integral to find the area of the ellipse

$$x = a \cos \theta, \quad y = b \sin \theta$$

Since the ellipse is regular with respect to both the axes, any one of the three results (5), (6) or (7) can be used.

$$\therefore \qquad \text{Required area} = -\int_C y \, dx = ab \int_0^{2\pi} \sin^2 \theta \, d\theta = \pi ab.$$

4.1 Deductions

Green's theorem has very important applications in Physics and Mathematics. We shall now make a few deductions which have very wide applications.

f, g, E etc. will all be supposed to satisfy the conditions of Green's theorem.

Deduction 1. *(Path independent line integral). If* $\dfrac{\partial g}{\partial x} = \dfrac{\partial f}{\partial y}$*, for every point of E, and if A and B are any two points of E, then the line integral* $\int (f\ dx + g\ dy)$ *has the same value for every path from A to B, provided the path lies in E.*

Let ACB and ADB be any two curves joining A and B, that lie in E and do not meet each other at any point other than A and B. Green's theorem holds for the region E' bounded by the curve $ACBDA$, but (by the given hypothesis) the double integral over E' is zero since $\partial g / \partial x = \partial f / \partial y$ at every point in E', and therefore the line integral along the bounding curve $ACBDA$ is zero.

$$0 = \int_{ACBDA} (f\ dx + g\ dy) = \int_{ACB} (f\ dx + g\ dy) + \int_{BDA} (f\ dx + g\ dy)$$

$$= \int_{ACB} (f\ dx + g\ dy) - \int_{ADB} (f\ dx + g\ dy)$$

$$\Rightarrow \qquad \int_{ACB} (f\ dx + g\ dy) = \int_{ADB} (f\ dx + g\ dy)$$

so that the two line integrals are equal and hence the result.

Deduction 2. *If the line integral* $\int_{AB} (f\ dx + g\ dy)$ *is independent of the path from A to B, where* ***A*** *and B are any two points of E, and the path AB lies in E, then* $\partial g / \partial x = \partial f / \partial y$*, for every point* (x, y) *of E.*

Let $E' \ (\subseteq E)$ be a region enclosed by a curve $ACBDA$ which lies in E, then

$$\iint_{E'} \left(\frac{\partial g}{\partial x} - \frac{\partial f}{\partial y} \right) dx\ dy = \int_{ACBDA} (f\ dx + g\ dy)$$

$$= \int_{ACB} (f\ dx + g\ dy) - \int_{ADB} (f\ dx + g\ dy) = 0$$

Now if the continuous function $(g_x - f_y)$ is not zero at an arbitrary point P of E, there is a region (neighbourhood) E' surrounding P in which $(g_x - f_y)$ has the same sign as at P and the integral over $ACBDA$ could not be zero, which is a contradiction. Hence $g_x - f_y = 0$ or $g_x = f_y$, at every point of E.

Deductions 1 and 2 may be stated in a combined form as:

If two functions f and g be continuous together with their partial derivatives f_y *and* g_x *on a bounded domain, piecewise regular with respect to both the axes, the line integral* $\int_C (f\ dx + g\ dy)$ *is independent of the path of integration lying within E if and only if* $\partial g / \partial x = \partial f / \partial y$*, for all points of E.*

Deduction 3. *If* $\partial g / \partial x = \partial f / \partial y$*, for every point of E, then* $(f\ dx + g\ dy)$ *is an exact differential.*

Let $P(\xi, \eta)$ be any point of E. It is always possible to choose another point $M(a, b)$ in E so that the path MNP where N is the point (ξ, b), lies in E. Let $F(\xi, \eta)$ be defined by the equation

$$F(\xi, \eta) = \int_{MNP} (f\ dx + g\ dy) + \text{constant}$$

$$= \int_a^{\xi} f(x, b)\ dx + \int_b^{\eta} g(\xi, y)\ dy + \text{constant}$$

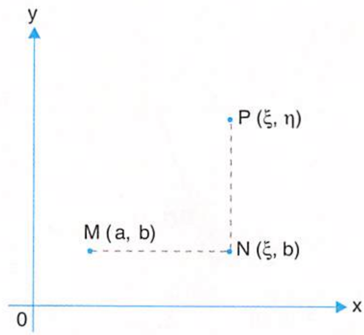


Fig. 14

Now

$$\frac{\partial F}{\partial \xi} = f(\xi,b) + \int_b^{\eta} \frac{\partial g(\xi,y)}{\partial \xi} dy = f(\xi,b) + \int_b^{\eta} \frac{\partial f(\xi,y)}{\partial y} dy$$

so that

$$\frac{\partial F}{\partial \xi} = f(\xi,b) + f(\xi,\eta) - f(\xi,b) = f(\xi,\eta)$$

Also $$\frac{\partial F}{\partial \eta} = g(\xi,\eta)$$

Hence

$$f(\xi,\eta)d\xi + g(\xi,\eta)d\eta = \frac{\partial F}{\partial \xi} d\xi + \frac{\partial F}{\partial \eta} d\eta = dF(\xi,\eta).$$

But, since (ξ, η) is an arbitrary point of E, therefore for every point (x, y) of E,

$$f\,dx + g\,dy = dF$$

Example 18. With the help of Green's formula, compute the difference between the line integrals

$$I_1 = \int_{ACB} \{(x + y)^2\ dx - (x - y)^2\ dy\}$$

and

$$I_2 = \int_{ADB} \{(x + y)^2\ dx - (x - y)^2\ dy\}$$

where ACB and ADB are respectively the straight line $y = x$, and the parabolic are $y = x^2$, joining the points $A(0, 0)$ and $B(1, 1)$.

■ Now,

$$I_2 - I_1 = \int_{ADBCA} \{(x + y)^2\, dx - (x - y)^2\, dy\}$$

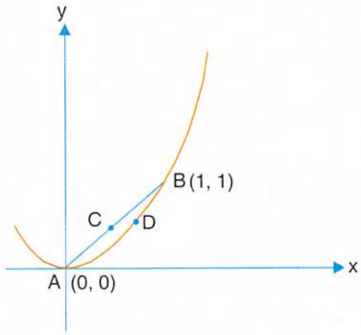


Fig. 15

The line integral is along the closed contour $ADBCA$ which encloses a domain E, bounded by $y = x$ and $y = x^2$ from (0, 0) to (1, 1) and quadratic with respect to y-axis (in fact both axes). Therefore by Green's theorem,

$$I_2 - I_1 = \iint_E \left\{ -\frac{\partial}{\partial x}(x - y)^2 - \frac{\partial}{\partial y}(x + y)^2 \right\} dx\, dy$$

$$= -4 \iint_E x\, dx\, dy$$

$$= -4 \int_0^1 x\, dx \int_{x^2}^{x} dy = -\frac{1}{3}$$

$\therefore$ $$|I_2 - I_1| = \frac{1}{3}$$

Example 19. Prove that the line integral

$$\int_C \frac{x\, dy - y\, dx}{x^2 + y^2}$$

taken in the positive direction over any closed contour with the origin inside it, is equal to 2π.

■ Since the origin is inside the closed contour, consider a circle (of contour Γ) with centre as origin and a convenient radius (say, unity) which does not intersect the given contour.

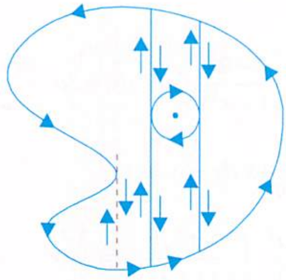


Fig. 16

Let E (with contour $C - \Gamma$) be the region enclosed by C and Γ. The region can be split up into domains quadratic with respect to the axes.

Hence by Green's theorem,

$$\int_{C-\Gamma} \frac{x\,dy - y\,dx}{x^2 + y^2} = \iint_E \left\{\frac{\partial}{\partial x}\left(\frac{x}{x^2 + y^2}\right) + \frac{\partial}{\partial y}\left(\frac{y}{x^2 + y^2}\right)\right\} dx\,dy$$

$$= \iint_E \left\{\frac{2}{x^2 + y^2} - \frac{2(x^2 + y^2)}{(x^2 + y^2)^2}\right\} dx\,dy = 0$$

$$\Rightarrow \qquad \int_C \frac{x\,dy - y\,dx}{x^2 + y^2} = \int_r \frac{x\,dy - y\,dx}{x^2 + y^2} = \int_0^{2\pi} d\theta = 2\pi\,.$$

Example 20. Using the line integral, compute the area of the loop of Descarte's folium $x^3 + y^3 = 3axy$.

■ To obtain the parametric equation of the contour of the folium, put $y = tx$, then

$$x = \frac{3at}{1 + t^3}, \quad y = \frac{3at^2}{1 + t^3}$$

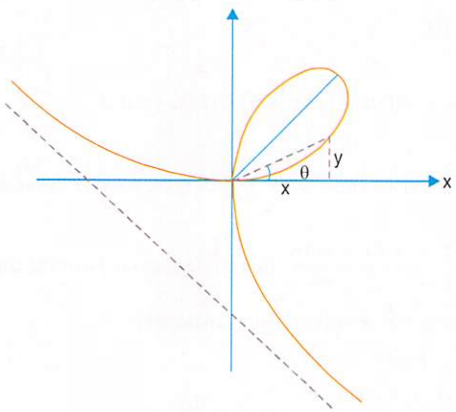


Fig. 17

From geometrical considerations, it is clear that the loop is described as t varies from 0 to ∞ (because $t = y/x = \tan\theta$ where θ varies from 0 to $\pi/2$).

We have

$$dx = 3a \frac{1 - 2t^3}{(1 + t^3)^2} dt, \quad dy = 3a \frac{2t - t^4}{(1 + t^3)^2} dt$$

$\therefore$ Area of the loop $= \frac{1}{2}\int_C (x\,dy - y\,dx)$

$$= \frac{9a^2}{2}\int_0^\infty \frac{t^2\,dt}{(1 + t^3)^2} = \frac{3}{2}a^2.$$

EXERCISE

1. Evaluate the following line integrals along C by two methods, (*i*) directly, (*ii*) as double integral by using Green's theorem:

   (*i*) $\int_C (1 - x^2)\,y\,dx + (1 + y^2)\,x\,dy$, where C is $x^2 + y^2 = a^2$

   (*ii*) $\int_C (xy + x + y)\,dx + (xy + x - y)\,dy$, where C is $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$

   (*iii*) Line integral in (*ii*), when C is the circle $x^2 + y^2 = ax$.

2. Prove that the line integral

$$\int_\Gamma \{(yx^3 + xe^y)\,dx + (xy^3 + ye^y - 2y)\,dy\}$$

   equals zero, if Γ is a closed curve symmetrical with respect to the origin.

3. Prove that the integral $\int_C (2xy - y)\,dx + x^2\,dy$, where C is a closed contour symmetrical with respect to the axes, is equal to the area of the domain bounded by C.

4. Show that the following integrals, taken around any closed contour C, are equal to zero.

   (*i*) $\int_C f\left(\frac{y}{x}\right)\frac{x\,dy - y\,dx}{x^2}$

   (*ii*) $\int_C [f(x + y) + f(x - y)]\,dx + [f(x + y) - f(x - y)]\,dy$.

5. Compute $\int_\Gamma \frac{x\,dy - y\,dx}{x^2 + 4y^2}$ round the circle $x^2 + y^2 = 1$ in the positive direction.

6. Evaluate the line integral, $\int_{(3,4)}^{(5,12)} \frac{x\,dx + y\,dy}{x^2 + y^2}$, (the origin does not lie on the contour.)

7. Using the line integral, compute the area of the figures enclosed by

   (*i*) Ellipse, $x = a\cos t$, $y = b\sin t$

   (*ii*) Asteroid, $x = a\cos^3 t$, $y = a\sin^3 t$

   (*iii*) Cardioid, $x = 2a\cos t - a\cos 2t$, $y = 2a\sin t - a\sin 2t$.

   (*iv*) Bernoulli's lemniscate, $(x^2 + y^2)^2 = 2a^2(x^2 - y^2)$.
   [*Hint*: Put $y = x\tan t$.]

ANSWERS

1. (*i*) $\frac{1}{2}\pi a^4$ (*ii*) 0 (*iii*) $-\pi a^3/8$

5. π 6. log 13/5

7. (*i*) πab (*ii*) $3\pi a^2/8$ (*iii*) $6\pi a^2$ (*iv*) $2a^2$

5. CHANGE OF VARIABLES

Lemma (*A preliminary formula*). Let u, v be the coordinates of a point of a region E_1 in the uv-plane, bounded by a curve (contour) C_1 and let x, y be the coordinates of a point of a region E bounded by a curve C in the xy-plane.

The transformation

$$x = X(u, v), \; y = Y(u, v) \qquad \text{...(1)}$$

where X and Y are two functions defined on the region E_1, establishes a correspondence between the points of the two regions. We suppose that

(*i*) the two functions X and Y possess continuous first order partial derivatives at all points of E_1 and C_1.

(*ii*) the equation (1) transforms the region E_1 with boundary C_1 into the region E with boundary C in such a way that a one-to-one correspondence exists between the two regions and their boundaries.

(*iii*) the Jacobian $\partial(X,Y)/\partial(u,v)$ does not change sign in E, though it may vanish at certain points of C_1.

When the point (u, v) describes the boundary C_1 in the positive sense (two cases arise) the point (x, y) may describe the boundary C in the positive or else in the negative sense without ever changing the sense of its motion. The transformation (correspondence) will be said to be *direct* or *inverse* respectively in the two cases.

We shall now obtain a formula connecting the area S of E with the area S' of E_1.

Let the curve C_1 be given by $u = u(t)$, $v = v(t)$ where t varies from a to b.

Consequently the curve C is given by

$$x = X(u(t), v(t)) = \phi(t)$$

$$y = Y(u(t), v(t)) = \Psi(t)$$

The area, $$S = \int_C x\, dy$$

taken along C in the positive sense, which may be expressed as

$$S = \int_a^b \phi \frac{d\Psi}{dt}\, dt = \int_a^b \phi\left[\frac{\partial Y}{\partial u}\frac{\partial u}{\partial t} + \frac{\partial Y}{\partial v}\frac{\partial v}{\partial t}\right] dt$$

Expressing this again as a line integral, we get

$$S = \pm \int_{C_1} X\left[\frac{\partial Y}{\partial u}du + \frac{\partial Y}{\partial v}dv\right] = \pm \int_{C_1} X(Y_u\ du + Y_v\ dv) \qquad \text{...(2)}$$

taken along C_1 in the positive sense, and the sign is to be taken + or – according as the transformation is direct or inverse.

By Green's theorem,

$$S = \pm \int_{C_1} X\,(Y_u\ du + Y_v\ dv) = \pm \iint_{E_1}\left[\frac{\partial}{\partial u}(X\ Y_v) - \frac{\partial}{\partial v}(X\ Y_u)\right] du\ dv$$

$$= \iint_{E_1} \frac{\partial(X,Y)}{\partial(u,v)} du\ dv$$

The partial derivatives Y_u, Y_v being continuous,

$$\therefore \qquad Y_{uv} = Y_{vu}$$

Again, using the mean value theorem (§ 2.7, VII), we get

$$S = \left[\frac{\partial(X,Y)}{\partial(u,v)}\right]_{(\xi,\eta)} \cdot \iint_{E_1} du\ dv = S'\cdot[J]_{(\xi,\eta)} \qquad \text{...(3)}$$

for some point (ξ, η) in E_1.

Since the areas S, S' are essentially positive, the sign + or – should be taken according as Jacobian J is positive or negative.

$$\therefore \qquad S = S' \mid J \mid_{(\xi,\,\eta)}$$

Note: If we compare the rule by which we selected the sign, + or –, in relations (2) and (3), we obtain the interesting property that

The transformation is direct or inverse according as the Jacobian J is positive or negative.

Main Theorem. Let a function f of two variables x, y be continuous on the region E. Let P_1 be a partition of the region E_1 and P the corresponding partition of E, induced by the transformation (1). Let $\Delta S_i'$ and ΔS_i be the areas of the corresponding sub-regions E_i' and E_i *of* E_1 and E respectively. Then

$$\Delta S_i = \Delta S' \mid J \mid_{(\xi_i,\eta_i)}$$

where (ξ_i, η_i) is some point of E_i'. Let (x_i, y_i) be the corresponding point of the sub-region E_i.

We have

$$f(x_i, y_i)\ \Delta S_i = f(X(\xi_i, \eta_i), Y(\xi_i, \eta_i)) \mid J \mid \Delta S_i'$$

Writing similar relation for each pair of the corresponding sub-regions, adding and letting the norm of the partitions tend to zero, we get

$$\iint_E f\,dx\,dy = \iint_E f(X, Y)\,|J|\,du\,dv.$$

ILLUSTRATIONS

1. To integrate $(x^2 + y^2)$ over the circle $x^2 + y^2 = a^2$, we change to polars,

$$x = r\cos\theta,\ y = r\sin\theta,$$

so that
$$|J| = \begin{vmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{vmatrix} = r$$

$$\iint_{x^2+y^2\le a^2} (x^2 + y^2)\,dx\,dy = \int_0^{2\pi}\int_0^a r^2 \cdot r\,dr\,d\theta = \frac{\pi a^4}{2}$$

2. To evaluate $\iint \sqrt{\dfrac{a^2b^2 - b^2x^2 - a^2y^2}{a^2b^2 + b^2x^2 + a^2y^2}}\,dx\,dy$, over the positive quadrant of the ellipse $x^2/a^2 + y^2/b^2 = 1$, by putting $x = au$, and $y = bv$, we have to evaluate

$$ab\iint \sqrt{\frac{1 - u^2 - v^2}{1 + u^2 + v^2}}\,du\,dv,$$ over the positive quadrant of the circle $u^2 + v^2 = 1$.

Again, changing to polars, by putting $u = r\cos\theta,\ v = r\sin\theta$, the integral becomes

$$ab\iint\sqrt{\frac{1 - r^2}{1 + r^2}}\,r\,dr\,d\theta$$

$$= ab\int_0^1 \sqrt{\frac{1 - r^2}{1 + r^2}}\,r\,dr \int_0^{\pi/2} d\theta$$

$$= \frac{1}{2}\pi ab \int_1^{\sqrt{2}} \sqrt{2 - \rho^2}\,d\rho, \text{ where } \rho^2 = 1 + r^2$$

$$= \frac{\pi}{2}\left(\frac{\pi}{4} - \frac{1}{2}\right)ab, \text{ taking } \rho = \sqrt{2}\sin t.$$

Note: The transformation could be effected in one step by putting $x/a = r\cos\theta,\ y/b = r\sin\theta$, then $|J| = abr$.

3. To evaluate $\iint \{2a^2 - 2a(x + y) - (x^2 + y^2)\}\,dx\,dy$ over the circle

$$x^2 + y^2 + 2a(x + y) = 2a^2,$$

transform the origin to $(-a, -a)$, by putting $x + a = u$, $y + a = v$, so that the integral becomes $\iint (4a^2 - u^2 - v^2)\,du\,dv$, over the circle $u^2 + v^2 = 4a^2$. Changing to polars, we get $8\pi a^4$.

CHAPTER 17

5.1 Solved Examples

Example 21. Evaluate $\iint (y - x)\, dx\, dy$, over the region E in the xy-plane bounded by the straight lines

$$y = x - 3,\ y = x + 1,\ 3y + x = 5,\ 3y + x = 7$$

- It is difficult to evaluate the double integral directly, however a simple change of coordinates reduces the domain of integration into a rectangle with sides parallel to the axes.

  Set

$$y - x = u,\ 3y + x = v$$

so that

$$x = \frac{1}{4}(v - 3u),\ y = \frac{1}{4}(v + u)$$

and

$$J = \frac{1}{16}\begin{vmatrix} -3 & 1 \\ 1 & 1 \end{vmatrix} = -\frac{1}{4} \quad \therefore\ |J| = -\frac{1}{4}.$$

The new domain is the rectangle $R = [-3, 1; 5, 7]$ in uv-plane.

$$\therefore \qquad \iint_E (y - x)\, dx\, dy = \iint_R u \cdot \frac{1}{4}\, du\, dv = \frac{1}{4}\int_{-3}^{1} u\, dv = -2$$

Example 22. Evaluate the integral

$$I = \int_0^1 dx \int_0^x \sqrt{x^2 + y^2}\, dy$$

by passing on to the polar coordinates.

- The integral in question is the double integral $\iint \sqrt{x^2 + y^2}\, dx\, dy$, over the region enclosed by the triangle $y = 0$, $y = x$, $x = 1$.

  In polar coordinates, the equations of these lines are $\theta = 0, \theta = \pi/4$, $r \cos\theta = 1$, so that the domain of integration is $0 \le \theta \le \pi/4$, $0 \le r \le \sec\theta$.

$$\therefore \qquad I = \int_0^{\pi/4} d\theta \int_0^{\sec\theta} r \cdot \bar{r}\, dr\, d\theta \; = \frac{1}{3}\int_0^{\pi/4} \sec^3\theta\, d\theta$$

$$= \frac{1}{6}\left[\sqrt{2} + \log\left(1 + \sqrt{2}\right)\right]$$

Example 23. Integrate $x^3 y^3$ over the area bounded by the parabolas

$$y^2 = ax,\ y^2 = bx,\ x^2 = py,\ x^2 = qy, \text{ where } 0 < a < b, \text{ and } 0 < p < q.$$

- The integration is to be carried over the shaded area, a curvilinear quadrangle, shown in the figure. Curves representing the opposite sides of this quadrangle are part of a *family of curves* which cover the entire xy plane and depend on one parameter. These two families can in fact be taken as the net

of curves, which partition the given region into sub-regions. These parameters usually give a convenient system of new coordinates for the given case.

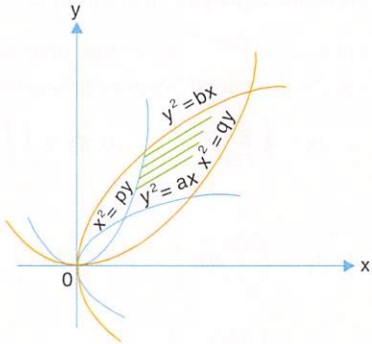


Fig. 18

Here the two families of parabolas are

$$y^2 = ux \quad (a \le u \le b)$$
$$x^2 = vy \quad (p \le v \le q)$$

Let us take u and v as the new coordinates, so that the new domain of integration is the rectangle $R = [a, b; p, q]$. Also $x = (uv^2)^{1/3}$, $y = (u^2 v)^{1/3}$, so that the Jacobian, $J = -\frac{1}{3}, |J| = \frac{1}{3}$.

Thus, the given integral

$$\iint x^3 y^3 \, dx \, dy = \iint_R u^3 v^3 \cdot \frac{1}{3} du \, dv$$
$$= \frac{1}{3} \int_a^b u^3 \, du \int_p^q v^3 \, dv = \frac{1}{48} (b^4 - a^4)(q^4 - p^4)$$

Example 24. Evaluate the *Euler-Poisson* integral

$$\int_0^\infty e^{-x^2} \, dx$$

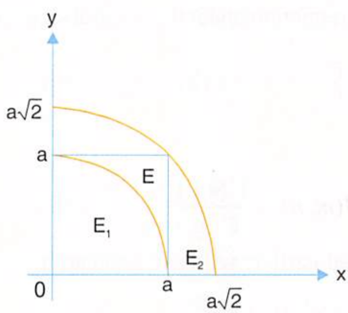


Fig. 19

■ The convergence of the integral was established earlier (Exp. 11(*iii*), Improper integrals); now we compute its value.

Let a region E be a square of side a and E_1, E_2 be quarter circles of radii $a, a\sqrt{2}$.

Let us consider an auxiliary function $e^{-(x^2+y^2)}$ over the three domains E_1, E and E_2. $e^{-(x^2+y^2)}$ being positive, its integral increases as the domain of integration is extended, hence

$$\iint_{E_1} e^{-(x^2+y^2)}\,dx\,dy \le \iint_{E} e^{-(x^2+y^2)}\,dx\,dy \le \iint_{E_2} e^{-(x^2+y^2)}\,dx\,dy \qquad ...(1)$$

Now,

$$\iint_{E_1} e^{-(x^2+y^2)}\,dx\,dy = \int_0^{\pi/2} d\theta \int_0^a re^{-r^2}\,dr = \frac{\pi}{4}(1-e^{-a^2})$$

$$\iint_{E_2} e^{-(x^2+y^2)}\,dx\,dy = \int_0^{\pi/2} d\theta \int_0^{a\sqrt{2}} re^{-r^2}\,dr = \frac{\pi}{4}(1-e^{-2a^2})$$

$$\iint_{E} e^{-(x^2+y^2)}\,dx\,dy = \int_0^a e^{-x^2}\,dx \int_0^a e^{-y^2}\,dy = \left[\int_0^a e^{-x^2}\,dx\right]^2$$

Hence (1) gives

$$\frac{\pi}{4}(1-e^{-a^2}) \le \left[\int_0^a e^{-x^2}\,dx\right]^2 \le \frac{\pi}{4}(1-e^{-2a^2})$$

As $a \to \infty$, the terms on the extreme ends tend to $\pi/4$.

$$\therefore \qquad \lim_{a\to\infty}\left[\int_0^a e^{-x^2}\,dx\right]^2 = \frac{\pi}{4}$$

Also, since the integrand is positive, we have

$$\int_0^\infty e^{-x^2}\,dx = \frac{\sqrt{\pi}}{2}$$

The integrand being an even function, we get

$$\int_{-\infty}^\infty e^{-x^2}\,dx = \sqrt{\pi}\,.$$

Ex. Assuming the validity of differentiation under the integral sign, prove that

$$\int_{-\infty}^\infty e^{-x^2}\cos(2xy)\,dx = \sqrt{\pi}\,e^{-y^2}$$

Example 25. Prove that

$$\beta(m,n) = \frac{\Gamma(m)\,\Gamma(n)}{\Gamma(m+n)},\ m>0,\ n>0$$

■ The convergence was discussed earlier; we now evaluate it.

Making the substitutions, $x=\cos^2\theta,\ t=x^2$

$$\beta(m, n) = \int_0^1 x^{m-1} (1-x)^{n-1}\, dx = 2 \int_0^{\pi/2} \cos^{2m-1}\theta \sin^{2n-1}\theta\, d\theta \qquad \text{...(1)}$$

$$\Gamma(m) = \int_0^\infty t^{m-1} e^{-t}\, dt = 2 \int_0^\infty x^{2m-1} e^{-x^2}\, dx \qquad \text{...(2)}$$

$$\therefore \qquad \Gamma(m)\,\Gamma(n) = 4 \int_0^\infty x^{2m-1} e^{-x^2}\, dx \int_0^\infty y^{2n-1} e^{-y^2}\, dy$$

$$= \lim_{a\to\infty} 4 \int_0^a x^{2m-1} e^{-x^2}\, dx \int_0^a y^{2n-1} e^{-y^2}\, dy$$

$$= \lim_{a\to\infty} 4 \iint_E x^{2m-1} y^{2n-1} e^{-(x^2+y^2)}\, dx\, dy$$

where E is a square of side a.

Taking quarter circle domains E_1 and E_2 and with the same reasoning as in the above example, we get

$$4 \iint_{E_1} x^{2m-1} y^{2n-1} e^{-(x^2+y^2)}\, dx\, dy \le 4 \iint_E x^{2m-1} y^{2n-1} e^{-(x^2+y^2)}\, dx\, dy$$

$$\le 4 \iint_{E_2} x^{2m-1} y^{2n-1} e^{-(x^2+y^2)}\, dx\, dy \qquad \text{...(3)}$$

Changing to polars,

$$4 \iint_{E_1} x^{2m-1} y^{2n-1} e^{-(x^2+y^2)}\, dx\, dy$$

$$= 4 \int_0^{\pi/2} \cos^{2m-1}\theta \sin^{2n-1}\theta\, d\theta \int_0^a r^{2(m+n)-1} e^{-r^2}\, dr$$

$$= 2\beta(m, n) \int_0^a r^{2(m+n)-1} e^{-r^2}\, dr$$

Similarly,

$$4 \iint_{E_2} x^{2m-1} y^{2n-1} e^{-(x^2+y^2)}\, dx\, dy = 2\beta(m, n) \int_0^{\sqrt{2}a} r^{2(m+n)-1} e^{-r^2}\, dr$$

Hence equation (3), gives

$$\beta(m, n) \int_0^a 2r^{2(m+n)-1} e^{-r^2}\, dr \le 4 \iint_E x^{2m-1} y^{2n-1} e^{-(x^2+y^2)}\, dx\, dy$$

$$\le \beta(m, n) \int_0^{\sqrt{2}a} 2r^{2(m+n)-1} e^{-r^2}\, dr$$

Proceeding to limits when $a \to \infty$, we get

$$\beta(m, n)\,\Gamma(m+n) \le \Gamma(m)\,\Gamma(n) \le \beta(m, n)\,\Gamma(m+n)$$

$$\Rightarrow \qquad \beta(m, n) = \frac{\Gamma(m)\,\Gamma(n)}{\Gamma(m+n)}$$

Example 26. Evaluate

$$\iint_E x^{m-1} y^{n-1} (1 - x - y)^{p-1} \, dx \, dy, \quad m \geq 1, n \geq 1, p \geq 1$$

where E is the region bounded by $x = 0$, $y = 0$, $x + y = 1$.

■ Now

$$\iint_E x^{m-1} y^{n-1} (1 - x - y)^{p-1} \, dx \, dy$$

$$= \int_0^1 dx \int_0^{1-x} x^{m-1} \, y^{n-1} \, (1 - x - y)^{p-1} \, dy$$

For this type of integrals, two sets of substitutions are possible, which give an integral with constant limits. We discuss both these here.

First method. Put $x + y = u$, $x = uv$ so that

$$y = u(1 - v)$$

and

$$J = \begin{vmatrix} v & u \\ 1 - v & -u \end{vmatrix} = -u$$

The Jacobian vanishes when $u = 0$, *i.e.*, when $x = 0 = y$. This origin of the xy-plane corresponds to the whole line $u = 0$ of the uv-plane, so that the correspondence ceases to be one-to-one. To exclude the origin of the xy-plane, we cut out the region along a line $x = h$ parallel to the y-axis and consider the integral on the remaining domain E_1, bounded by the lines

$$v = 1, \; u = 1, \; uv = h$$

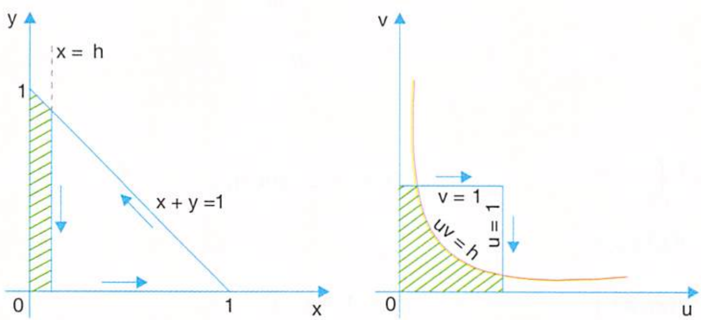


Fig. 20

However, in the limit when $h \to 0$, this new region degenerates into the square bounded by

$$v = 1, \; u = 1, \; v = 0, \; u = 0$$

Thus,

$$\iint_E x^{m-1} y^{n-1} (1 - x - y)^{p-1} \, dx \, dy$$

$$= \int_0^1 \int_0^1 u^{m-1} v^{m-1} u^{n-1} (1-v)^{n-1} (1-u)^{p-1}\, u\, du\, dv$$

$$= \int_0^1 u^{m+n-1} (1-u)^{p-1}\, du \int_0^1 v^{m-1} (1-v)^{n-1}\, dv$$

$$= \beta(m+n, p) \cdot \beta(m, n)$$

$$= \frac{\Gamma(m+n)\,\Gamma(p)}{\Gamma(m+n+p)} \frac{\Gamma(m)\,\Gamma(n)}{\Gamma(m+n)} = \frac{\Gamma(m)\,\Gamma(n)\,\Gamma(p)}{\Gamma(m+n+p)}$$

Second method. Put $x = u,\ y = (1-u)v$, so that

$$1 - x - y = 1 - u - (1-u)\, v = (1-u)(1-v)$$

and

$$J = \begin{vmatrix} 1 & 0 \\ -v & 1-u \end{vmatrix} = 1 - u$$

which vanishes for $u = 1$, so that the point $(1, 0)$ in the xy-plane corresponds to the whole line $u = 1$ in the uv-plane. However, proceeding as in the first method, the new region becomes the square

$$u = 0,\ v = 0,\ u = 1,\ v = 1$$

Thus

$$\iint_E x^{m-1} y^{n-1} (1-x-y)^{p-1}\, dx\, dy$$

$$= \int_0^1 \int_0^1 u^{m-1} v^{n-1} (1-u)^{p+n-1} (1-v)^{p-1}\, du\, dv$$

$$= \int_0^1 u^{m-1} (1-u)^{p+n-1}\, du \int_0^1 v^{n-1} (1-v)^{p-1}\, dv$$

$$= \beta(m, p+n) \cdot \beta(n, p)$$

$$= \frac{\Gamma(m)\,\Gamma(p+n)}{\Gamma(m+n+p)} \cdot \frac{\Gamma(n)\,\Gamma(p)}{\Gamma(p+n)} = \frac{\Gamma(m)\,\Gamma(n)\,\Gamma(p)}{\Gamma(m+n+p)}.$$

Ex. Show that

(*i*) $\iint_E \sqrt{xy}\, dx\, dy = \pi/24$,

(*ii*) $\iint_E \sqrt{xy\,(1-x-y)}\, dx\, dy = 2\pi/105$,

where E is the region bounded by the lines $x = 0, y = 0, x + y - 1 = 0$.

Example 27. Evaluate $\int_0^\pi \int_0^\pi |\cos(x+y)|\, dx\, dy$.

■ Put $x = u - v, y = v$, so that $x + y = u$, and $J = 1$

$$\therefore \quad \int_0^\pi \int_0^\pi |\cos(x+y)|\, dx\, dy = \int_0^\pi dv \int_v^{v+\pi} |\cos u|\, du$$

CHAPTER 17

Now

$$\int_v^{v+\pi} |\cos u| \, du = \int_v^{\pi/2} |\cos u| \, du + \int_{\pi/2}^{\pi} |\cos u| \, du + \int_{\pi}^{\pi+v} |\cos u| \, du$$

$$= \sin u \Big|_v^{\pi/2} + (-\sin u) \Big|_{\pi/2}^{\pi} + (-\sin u) \Big|_{\pi}^{\pi+v}$$

$$= 1 - \sin v + 1 - \sin(\pi + v) = 2$$

Hence

$$\int_0^{\pi} \int_0^{\pi} |\cos(x + y)| \, dx \, dy = \int_0^{\pi} 2 dv = 2\pi.$$

Example 28. Evaluate $\iint_E \sin\left(\frac{x-y}{x+y}\right) dx \, dy$, where E is the region bounded by the co-ordinate axes and $x + y = 1$ in the first quadrant.

■ Taking $x - y = u$, $x + y = v$, so that $x = \frac{1}{2}(u + v)$, $y = \frac{1}{2}(v - u)$, and the Jacobian is $\frac{1}{2}$.

$$\therefore \quad \iint_E \sin\left(\frac{x-y}{x+y}\right) dx \, dy = \iint_{E_{uv}} \sin\left(\frac{u}{v}\right) \frac{1}{2} du \, dv \qquad ...(1)$$

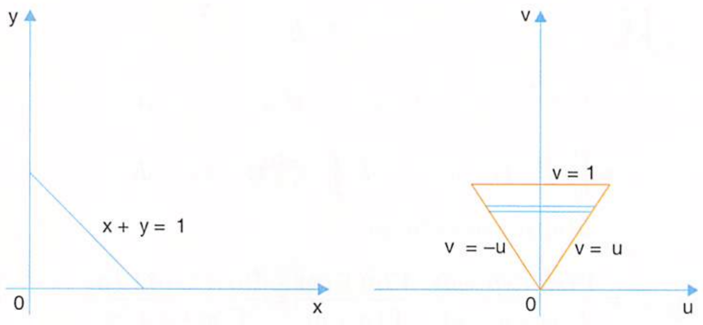


Fig. 21

Now

$$\iint_{E_{uv}} \sin\left(\frac{u}{v}\right) \frac{1}{2} du \, dv = \frac{1}{2} \int_0^1 dv \int_{-v}^{v} \sin \frac{u}{v} du$$

$$= \frac{1}{2} \int_0^1 v\{-\cos 1 + \cos(-1)\} \, dv = 0 \qquad ...(2)$$

Hence from (1) and (2), the required integral is zero.

Ex. 1. Prove that $\iint_E \exp\{(y - x)/(y + x)\} \, dx \, dy = (e^2 - 1)/4e$, where

$$E = \{(x, y) \mid x \geq 0, y \geq 0, x + y \leq 1\}.$$

Ex. 2. Prove that $\iint_E \sin x \sin y \sin(x+y)\, dx\, dy = \pi/16$, where

$$E = \{(x, y) \mid x \geq 0, y \geq 0, x + y \leq \pi/2\}.$$

5.2 Areas of Plane Regions

It was shown in § 3.2 that the area S of a plane region E in the xy-plane, bounded by a closed curve, is given by

$$S = \iint_E dx\, dy$$

In polar coordinates, it can be expressed as

$$S = \iint_E r\, dr\, d\theta.$$

ILLUSTRATIONS

1. Area S enclosed by the circle $x^2 + y^2 = a^2$ is given by

$$S = \int_{-a}^{a} dx \int_{-\sqrt{(a2-x2)}}^{\sqrt{(a2-x2)}} dy.$$

Using the symmetry of the region, we get

$$S = 4\int_0^a dx \int_0^{\sqrt{(a2-x2)}} dy = 4\int_0^{\pi/2} d\theta \int_0^a r\, dr = \pi a^2.$$

2. Area enclosed by $r = 1 - \cos\theta$ is

$$\int_0^{2\pi} d\theta \int_0^{1-\cos\theta} r\, dr = \frac{3}{2}\pi.$$

EXERCISE

1. Compute the double integral by passing over to polar coordinates:

   (i) $\iint_E \sqrt{a^2 - x^2 - y^2}\, dx\, dy$, where E is the region bounded by the circle $x^2 + y^2 = ax$

   (ii) $\int_0^a dx \int_0^{\sqrt{(a2-x2)}} \log(1 + x^2 + y^2)\, dy$.

2. Find the areas of the domains enclosed by

   (i) $y = x, y = 5x, x = 1$

   (ii) $y = \sqrt{x}, y = 2\sqrt{x}, x = 4$

   (iii) $(x^2 + y^2)^2 = 2ax^3$

   (iv) The loop of $(x + y)^3 = xy$ lying in the first quadrant

   (v) $(x^2 + y^2)^2 = 2a^2(x^2 - y^2)$.

CHAPTER 17

3. Evaluate the integral

$$\int_0^{2a} dx \int_{\sqrt{(2ax-x2)}}^{\sqrt{(4ax-x2)}} \left(1+\frac{y^2}{x^2}\right) dy$$

by changing the coordinates to $r,\ \theta$ where

$$x = r\cos^2\theta,\ \ y = r\sin\theta\cos\theta\cdot$$

4. Use the transformation

$$u = (x^2+y^2)/x,\quad v = (x^2+y^2)/y\cdot$$

to evaluate the integral $\iint xy\,dx\,dy$, over the region common to the circles $x^2+y^2=x,\ x^2+y^2=y$.

5. Evaluate $\iint_E \dfrac{y+1}{x^2+(y+1)^2}\,dx\,dy$, where

$$E = \{(x, y) \mid y \ge 0,\ x^2+y^2 \le 1\}.$$

6. Evaluate $\iint_E (x+y)^n\,dx\,dy$, n being positive, and

$$E = \{(x, y) \mid x \ge 0,\ y \ge 0,\ x+y \le 1\}.$$

7. Evaluate $\iint_E \left(x^2+y^2\right)dx\,dy$, where E is the area of the xy-plane defined by the relations:

$$2a \ge x^2+y^2 \ge a,\ x \ge 0,\ a \ge x^2-y^2 \ge -a,\ y \ge 0$$

by means of the transformation

$$x^2+y^2=u,\ \ x^2-y^2=v.$$

8. Show that

$$\iint \sqrt{(4a^2-x^2-y^2)}\ dx\ dy = \frac{4}{9}(3\pi-4)a^3.$$

taken over the upper half of the circle $x^2+y^2-2ax=0$.

9. Show that

$$\iint_E \frac{dx\,dy}{xy} = \log\frac{a'}{a}\log\frac{b'}{b}$$

where E is the area bounded by the four circles

$$x^2+y^2 = ax,\ a'x,\ by,\ b'y$$

$a,\ a',\ b,\ b'$ being all positive.

10. By changing to polar coordinates, show that

$$\iint_E \sqrt{x^2+y^2}\ dx\ dy = \frac{38\pi}{3}$$

where E is the region in the xy-plane bounded by

$$x^2+y^2=4,\ \ x^2+y^2=9$$

11. Using the transformation $x+y=u$, $y=uv$, show that

$$\int_0^1 dx \int_0^{1-x} e^{y/(x+y)}\ dy = \frac{1}{2}(e-1)$$

12. When the field of integration is $y = 0, y = x, x = 1$, show that

$$\iint \sqrt{(4x^2 - y^2)}\, dx\, dy = \frac{1}{3}\left(\frac{\pi}{3} + \frac{\sqrt{3}}{2}\right).$$

13. Prove that

$$\iint_E x^{m-1} y^{n-1} f(x+y)\, dx\, dy = \frac{\Gamma(m)\,\Gamma(n)}{\Gamma(m+n)} \int_0^1 t^{m+n-1} f(t)\, dt$$

where $E = \{(x, y) : x, y \geq 0,\ x + y \leq 1\}$ and m, n are less than unity.

14. Prove that, for $l, m, n \geq 1$,

$$\int_0^1 y^{m-1}\, dy \int_0^{1-y} x^{2l-1}(1+x)^{m+n-1}(1-x-y)^{n-1}\, dx = \frac{1}{2}\frac{\Gamma(l)\Gamma(m)\Gamma(n)}{\Gamma(l+m+(n)}$$

15. If $m \geq 0$, prove that the integral

$$\iint \left(1 - \frac{x^2}{a^2} - \frac{y^2}{b^2}\right)^m f(Ax + By)\, dx\, dy$$

taken over the ellipse $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = 1$ is equal to

$$\beta\left(\tfrac{1}{2}, m+1\right) ab \int_{-1}^{1} (1 - x^2)^{m+\frac{1}{2}} f(kx)\, dx,\ \ k = \sqrt{A^2a^2 + B^2b^2}\,.$$

16. Show that

$$\int_0^{\pi/2} d\phi \int_0^{\pi/2} f(1 - \sin\theta\cos\phi)\sin\theta\, d\theta = \frac{\pi}{2}\int_0^1 f(x)\, dx\,.$$

ANSWERS

1. (*i*) $\dfrac{a^3}{3}\left(\pi - \dfrac{4}{3}\right)$, (*ii*) $\dfrac{\pi}{4}\{(1+a^2)\log(1+a^2) - a^2\}$

2. (*i*) 2, (*ii*) 16/3, (*iii*) $5\pi a^2/8$, (*iv*) 1/60, (*v*) $2a^2$

3. $(3\pi + 8)a^2/3$ 4. 1/96 5. 1 6. $1/6(n+4)$

7. $\pi a^2/24 + \sqrt{3}a^2/4$

CHAPTER 17

18 Integration on $\mathbf{R}^3$

(Line, Surface and Volume Integrals)

This chapter deals mainly with three types of integrals in a three-dimensional space, viz.,

(*i*) Line integrals,

(*ii*) Surface integrals,

(*iii*) Volume (triple) integrals.

The discussion starts with the consideration of rectifiable curves and rectification, and goes on to consider the line integrals, surface areas and the surface integrals, the volumes and the volume (triple) integrals in that order. Stokes' theorem, which connects a line integral with a double integral, and Gauss's theorem, which establishes a relation between a surface integral and a volume integral, have also been considered.

1. RECTIFIABLE CURVES

We are now in a position to consider some suitable definitions of the length of the curve (rectification) and to discuss the conditions under which a curve is rectifiable. *We shall be concerned with continuous curves only* in the present discussion. The reader will see the importance of the functions of bounded variation in such a discussion.

A *curve* in space is defined to be a vector-valued function γ with domain as a subset of $\mathbf{R}$ and range a subset of $\mathbf{R}^3$. The curve is *continuous* if γ is continuous, and is called a *Jordan arc* if γ is one-to-one.

1.1 Length of a Curve

Let $\gamma = (X, Y, Z)$, where X, Y, Z are three real *single-valued, continuous* functions of t defined on an interval $[a, b]$, be a *continuous curve* in $\mathbf{R}^3$, so that $\gamma(t) = [X(t), Y(t), Z(t)]$ is a point on the curve* corresponding to $t \in [a, b]$.

Corresponding to any partition

$$P = \{a = t_0, t_1, t_2, \ldots, t_n = b\}$$

of $[a, b]$, we get an ordered set of points

$$\{\gamma(t_0), \gamma(t_1), \ldots, \gamma(t_n)\}$$

---

* $x = X(t), y = Y(t), z = Z(t), \ t \in [a, b]$

may be thought of as a (parametric) representation of the curve γ in $\mathbf{R}^3$.

on the curve γ. This set of points may be thought of as a polygon inscribed in the curve.

$|\gamma(t_i) - \gamma(t_{i-1})|$ represents the distance between the points $\gamma(t_{i-1})$ and $\gamma(t_i)$ and so the sum

$$s(P, a, b) = \sum_{i=1}^{n} |\gamma(t_i) - \gamma(t_{i-1})|$$

is the length of the polygon inscribed in the curve with vertices at the points $\gamma(t_0), \gamma(t_1), \ldots, \gamma(t_n)$ in this order. The length of the polygon clearly depends upon the particular partition of $[a, b]$. Let us consider the set of these sums for all possible partitions of $[a, b]$.

As the mesh of the partition tends to zero (or $n \to \infty$) these polygons approach the curve more and more closely. Accordingly,

The curve is said to be *rectifiable* if this set of sums is bounded and then, the supremum of the set is known as the *length of the curve.*

Thus a continuous curve γ is *rectifiable* when the supremum of the sums $\Sigma|\gamma(t_i) - \gamma(t_{i-1})|$, taken over all partitions of $[a, \mathrm{b}]$, exists, *i.e.*, when γ is of bounded variation over $[a, b]$. The length, $s(a, b)$ of the curve, from $t = a$ to $t = b$ is then the total variation of γ over $[a, b]$. Thus

$$s(a, b) = V(\gamma, a, b)$$

The length of the curve over the subinterval $[a, t]$ is generally denoted by $s(t)$, where

$$s(t) = s(a, t) = V(\gamma, a, t) = v_\gamma(t)$$

is a function of t.

Note: $\gamma = (\phi, \Psi, \theta)$ being a vector-valued function, the meaning of $|\gamma(t_i) - \gamma(t_{i-1})|$ or $V(\gamma, a, b)$ should be clearly understood. Here,

$$|\gamma(t_i) - \gamma(t_{i-1})| = \sqrt{\{[\phi(t_i) - \phi(t_{i-1})]^2 + [\Psi(t_i) - \Psi(t_{i-1})]^2 + [\theta(t_i) - \theta(t_{i-1})]^2\}}\,.$$

1.2 Properties of Rectifiable Curves

We know that a rectifiable continuous curve in space is a continuous vector-valued function (domain $\mathbf{R}$ and range $\mathbf{R}^3$) of bounded variation. The following properties of such curves are the direct consequences of the vector-valued functions of bounded variation.

1. Since for a vector-valued function, $\gamma = (X, Y, Z)$, of bounded variation,

$$V(X, a, b) \le V(\gamma, a, b) \le V(X, a, b) + V(Y, a, b) + V(Z, a, b)$$

we at once deduce that

$$V(X, a, b) \le s(a, b);\ V(Y, a, b) \le s(a, b);\ V(Z, a, b) \le s(a, b)$$

and

$$s(a, b) \le V(X, a, b) + V(Y, a, b) + V(Z, a, b)\,.$$

2. A curve $\gamma = (X, Y, Z)$ is rectifiable if X, Y, Z are derivable and the derivatives are bounded.

Under the given hypothesis, γ' exists and is bounded and therefore γ is of bounded variation. Consequently, γ is rectifiable.

CHAPTER 18

3. Since a function of bounded variation over $[a, b]$ is of bounded variation over each sub-interval of $[a, b]$, an arc (part) of a rectifiable curve corresponding to any of its sub-intervals is also rectifiable.
4. Since for a function of bounded variation,
$$V(\gamma, a, b) = V(\gamma, a, c) + V(\gamma, c, b), \quad a \le c \le b$$
it follows that
$$s(a, b) = s(a, c) + s(c, b)$$
5. The arc $s(t)$ of a rectifiable curve is a monotone increasing function of t.
Since the total variation function $v_\gamma(t)$ of a function of bounded variation is monotone increasing over $[a, b]$, the arc $s(t)$ is also monotone increasing on $[a, b]$.
Again since $v_\gamma(t)$ is a *strictly* monotone increasing function over $[a, b]$ unless X, Y, Z (or equivalently γ) are constant functions over some sub-interval of $[a, b]$, it follows that the arc $s(t)$ is also a strictly monotone increasing function unless all the three functions X, Y, Z (or γ) are constant over that sub-interval.
6. The arc $s(t)$ is continuous over $[a, b]$.
Since we are dealing with continuous curves, X, Y, Z and therefore γ are all continuous. Also, since the total variation function $v_\gamma(t)$ of a continuous function is continuous, the arc $s(t)$ is continuous over $[a, b]$.

Remark: If γ (or equivalently, X, Y, Z) is not constant over any sub-interval of $[a, b]$, $s(t)$ is a strictly monotone and increasing continuous function of t on $[a, b]$, therefore the function s is invertible and s^{-1} is continuous on $[a, b]$.

Thus t can be regarded as a strictly monotone increasing function of s over $[s(a), s(b)]$, so that the arc length s may be used as a parameter in place of t.

1.3 The Riemann Integral for Length of a Curve

In certain cases, the length of a curve can be found with the help of a Riemann integral. We shall prove this for *smooth* curves, *i.e.*, for curves which have no multiple points and which are continuously differentiable. Thus if $\gamma = (X, Y, Z)$ is a continuously differentiable curve on $[a, b]$, then X', Y', Z' all exist, are continuous and do not vanish simultaneously on $[a, b]$. The last condition is equivalent to saying that $(X'^2 + Y'^2 + Z'^2)$ does not vanish for any value of t on $[a, b]$.

Theorem 1. *If γ is a smooth curve in R^3 such that γ' exists and is continuous on $[a, b]$, then γ is rectifiable and has a length*
$$\int_a^b |\gamma'(t)|\, dt$$

Since the function γ has a bounded derivative, it is of bounded variation on $[a, b]$ and is therefore rectifiable on $[a, b]$. We have to prove that,

$$\int_a^b |\gamma'(t)|\, dt = V(\gamma, a, b)$$

For any partition $\{a = t_0, t_1, t_2, ..., t_n = b\}$ of $[a, b]$, we have

$$\Sigma|\gamma(t_i) - \gamma(t_{i-1})| = \Sigma\left|\int_{t_{i-1}}^{t_i} \gamma'(u)\, du\right| \le \Sigma\int_{t_{i-1}}^{t_i} |\gamma'(u)|\, du = \int_a^b |\gamma'(t)|\, dt$$

$$\Rightarrow \qquad V(\gamma, a, b) \le \int_a^b |\gamma'(t)|\, dt \qquad ...(1)$$

Let $\varepsilon > 0$ be given.

Since γ' is uniformly continuous on $[a, b]$, there exists $\delta > 0$ such that

$$|\gamma'(v) - \gamma'(u)| < \varepsilon, \text{ if } |v - u| < \delta$$

Let $P = \{a = t_0, t_1, ..., t_n = b\}$ be a partition of $[a, b]$ with mesh $\mu(P) < \delta$.

Now for $t_{i-1} \le u \le t_i$,

$$|\gamma'(t_i) - \gamma'(u)| < \varepsilon$$

or

$$|\gamma'(u)| - \varepsilon < |\gamma'(t_i)|$$

so that, integrating from t_{i-1} to t_i, we get

$$\int_{t_{i-1}}^{t_i} |\gamma'(u)|\, du - \varepsilon\, \Delta t_i < |\gamma'(t_i)|\, \Delta t_i - |\gamma'(t_i)\, \Delta t_i|$$

$$= \left|\int_{t_{i-1}}^{t_i} [\gamma'(u) + \gamma'(t_i) - \gamma'(u)]\, du\right|$$

$$\le \left|\int_{t_{i-1}}^{t_i} \gamma'(u)\, du\right| + \left|\int_{t_{i-1}}^{t_i} [\gamma'(t_i) - \gamma'(u)]\, du\right|$$

$$\le |\gamma(t_i) - \gamma(t_{i-1})| + \int_{t_{i-1}}^{t_i} |\gamma'(t_i) - \gamma'(u)|\, du$$

$$\Rightarrow \qquad \int_{t_{i-1}}^{t_i} |\gamma'(u)|\, du < |\gamma(t_i) - \gamma(t_{i-1})| + 2\varepsilon\, \Delta t_i$$

Writing these inequalities for $i = 1, 2, ..., n$, and adding, we get

$$\int_a^b |\gamma'(t)|\, dt < \sum_{i=1}^{n} |\gamma(t_i) - \gamma(t_{i-1})| + 2\varepsilon(b - a)$$

$$\le V(\gamma, a, b) + 2\varepsilon(b - a)$$

But since ε is arbitrary,

$$\int_a^b |\gamma'(t)\, dt \le V(\gamma, a, b) \qquad ...(2)$$

(1) and (2) imply

$$\int_a^b |\gamma'(t)|\, dt = V(\gamma, a, b)$$

i.e., the length of the curve is $\int_a^b |\gamma'(t)|\, dt$.

Remark: The theorem may be stated as follows:

If $\gamma = (X, Y, Z)$ is a curve such that X', Y', Z' exist, and are continuous on $[a, b]$ and do not vanish simultaneously for any $t \in [a, b]$, then the curve is rectifiable and has a length $\int_a^b |\gamma'(t)|\, dt$.

Note: Since γ is a vector-valued function,

$$|\gamma'(t)| = \sqrt{X'^2(t) + Y'^2(t) + Z'^2(t)}\,.$$

Example 1. Find the length of the curve

$$x = at^2,\ y = 2at,\ z = at,\ 0 \le t \le 1$$

■
$$s(0, 1) = \int_0^1 \sqrt{(2at)^2 + (2a)^2 + a^2}\ dt = a\int_0^1 \sqrt{5 + 4t^2}\ dt$$

Put $2t = \sqrt{5}\sinh u$.

$$s(0, 1) = \frac{5a}{4}\int_0^{\log\sqrt{5}} (1 + \cosh 2u)\, du = \frac{5a}{4}\left[\log\sqrt{5} + \frac{6}{5}\right] = \frac{a}{8}\,[5\log 5 + 12].$$

Ex. Show that the length of the curve

1. $x = a\cos\theta,\ y = a\sin\theta,\ z = a\theta,\ 0 \le \theta \le 2\pi$ is $2\sqrt{2}a\pi$.
2. $x = 2t - 1,\ y = t + 1,\ z = t - 2,\ 0 \le t \le 3$ is $3\sqrt{6}$.
3. $x = a(\theta - \sin\theta),\ y = a(1 - \cos\theta),\ z = a\theta,\ -\pi \le \theta \le \pi$ is

$$\int_0^{\pi} 2a\sqrt{3 - 2\cos\theta\, d\theta}\,\cdot$$

4. $x = 2t,\ y = 2t + 1,\ z = t^2 + t$ between the points (0, 1, 0) and (2, 3, 2) is

$$3\left(\sqrt{17 - 1}\right)/4 + \log\left\{\left(13 + 3\sqrt{17}\right)/8\right\}.$$

2. LINE INTEGRALS

We have considered line integrals of functions along plane curves. In this section we shall consider line integrals along space curves. Most of the facts of the theory of plane curves are automatically generalized to the case of space curves.

Definition. Let $x = X(t),\ y = Y(t),\ z = Z(t),\ a \le t \le b$
be a curve C in a space of three dimensions.

Let a bounded vector-valued function $F = (f, g, h)$ be defined at every point of the curve C, so that f, g, h are all bounded and defined at every point (x, y, z) of the curve.

Let $\{a = t_0, t_1 \ldots, t_n = b\}$ be any partition of $[a, b]$, and ξ_i any point of Δt_i. Let (x_i, y_i, z_i) be a point of the curve corresponding to $t = t_i$, where

$$x_i = X(t_i),\ \ y_i = Y(t_i),\ \ z_i = Z(t_i)$$

Form the sum

$$\sum_{i=1}^{n} [f(X(\xi_i), Y(\xi_i), Z(\xi_i))\, \Delta x_i + g(X(\xi_i), Y(\xi_i), Z(\xi_i))\, \Delta y_i$$
$$+ h(X(\xi_i), Y(\xi_i), Z(\xi_i))\, \Delta z_i]$$

If, as the norm of the partition tends to zero, the sum tends to a finite limit, independent of the choice of ξ_i in Δt_i, the line integral of F along C exists and is denoted by $\int_C F\, dt$ or

$$\left.\begin{aligned} &\int_C f(X, Y, Z)\, dx + g(X, Y, Z)\, dy + h(X, Y, Z)\, dz \\ \text{or}\quad &\int_C f\, dx + g\, dy + h\, dz \end{aligned}\right\} \quad \ldots(1)$$

and call it the *line integral* of $F = (f, g, h)$ along C.

2.1 A Sufficient Condition for Existence

Applying arguments completely similar to those used for the case of a line integral along plane curves (§ 1.4, Ch. 16), it can be proved that if

(*i*) X, Y, Z possess continuous derivatives in $[a, b]$, and

(*ii*) $F = (f, g, h)$ is continuous at every point of C,

then the integral $\int_C f\, dx + g\, dy + h\, dz$ exists and is equal to the ordinary integral

$$\int_a^b \{f(X, Y, Z)X' + g(X, Y, Z)Y' + h(X, Y, Z)Z'\}\, dt\,.$$

Note: For evaluating a line integral, the parametric equation of the curve should be known.

Remark: *Vectorial formulation*. Let $\mathbf{r}$ be the position vector of a point (x, y, z) on the curve, so that

$$\mathbf{r} = \mathbf{i}x + \mathbf{j}y + \mathbf{k}z$$
$$= \mathbf{i}X(t) + \mathbf{j}Y(t) + \mathbf{k}Z(t),\ \ a \le t \le b$$

where $\mathbf{i}, \mathbf{j}, \mathbf{k}$ are the unit vectors along the coordinate axes.

Let a vector function be represented as

$\mathbf{F}(x, y, z) = \mathbf{i}P(x, y, z) + \mathbf{j}Q(x, y, z) + \mathbf{k}R(x, y, z)$

$$\therefore \qquad \mathbf{F}\cdot\frac{d\mathbf{r}}{dt} = P\frac{dx}{dt} + Q\frac{dy}{dt} + R\frac{dz}{dt}$$

$$\Rightarrow \qquad \int_a^b \left(\mathbf{F}\cdot\frac{d\mathbf{r}}{dt}\right) dt = \int_\alpha^\beta \left(P\frac{dx}{dt} + Q\frac{dy}{dt} + R\frac{dz}{dt}\right) dt$$

Thus the line integral can be expressed as

$$\int_C \mathbf{F}\cdot d\mathbf{r} = \int_a^b \left(\mathbf{F}\cdot\frac{d\mathbf{r}}{dt}\right) dt$$

2.2 Line Integral with Respect to Arc Length

We have seen that the line integral, $\int f\,dx + g\,dy + h\,dz$ of a function $F = (f, g, h)$ along a curve C,

$$x = X(t),\quad y = Y(t),\quad z = Z(t),\quad a \le t \le b$$

can be expressed as an ordinary integral

$$\int_a^b \{f(X, Y, Z)X' + g(X, Y, Z)Y' + h(X, Y, Z)Z'\}\,dt$$

If the curve is *smooth*, it can be represented as

$$x = \theta(s),\quad y = \phi(s),\quad z = \Psi(s)$$

where the arc length s varies from 0 to l (l being the length of the curve as t varies from a to b).

Considering the partitions $\{0 = s_0, s_1 \ldots, s_n = l\}$ and proceeding as in § 2.1 (with parameter s in place of t) we deduce that the line integral

$$\int_C f\,dx + g\,dy + h\,dz \text{ reduces to}$$

$$\int_0^l \left\{f(\theta, \phi, \Psi)\frac{dx}{ds} + g(\theta, \phi, \Psi)\frac{dy}{ds} + h(\theta, \phi, \Psi)\frac{dz}{ds}\right\} ds$$

which is equivalent to the ordinary integral, or

$$\int_a^b \left(f\frac{dx}{ds} + g\frac{dy}{ds} + h\frac{dz}{ds}\right)\frac{ds}{dt}\,dt$$

where $$\frac{ds}{dt} = \sqrt{\left(\frac{dX}{dt}\right)^2 + \left(\frac{dY}{dt}\right)^2 + \left(\frac{dZ}{dt}\right)^2}.$$

Example 2. Prove that

$$\int \frac{x^2 + y^2}{p}\,ds = \frac{\pi ab}{4}[4 + (a^2 + b^2)(a^{-2} + b^{-2})]$$

when the integral is taken round the ellipse $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = 1$, and p is the length of the perpendicular from the centre to the tangent.

- The parametric equation of the ellipse is

$$x = a \cos \phi, \quad y = b \sin \phi$$

$$p = \frac{y - x(dy/dx)}{\sqrt{1 + (dy/dx)^2}} = \frac{ab}{\sqrt{a^2 \sin^2 \phi + b^2 \cos^2 \phi}}$$

$$\frac{ds}{d\phi} = \sqrt{a^2 \sin^2 \phi + b^2 \cos^2 \phi}$$

$$\therefore \quad \int_C \frac{x^2 + y^2}{p} \, ds = \int_0^{2\pi} \frac{x^2 + y^2}{p} \frac{ds}{d\phi} \, d\phi$$

$$= \frac{1}{ab} \int_0^{2\pi} (a^2 \cos^2 \phi + b^2 \sin^2 \phi)(a^2 \sin^2 \phi + b^2 \cos^2 \phi) \, d\phi$$

$$- \frac{4}{ab} \int_0^{\pi/2} [(a^4 + b^4) \cos^2 \phi \sin^2 \phi + a^2 b^2 (\cos^4 \phi + \sin^4 \phi)] \, d\phi$$

$$= \frac{\pi}{4ab} [a^4 + b^4 + 6a^2 b^2] = \frac{\pi ab}{4} [(a^2 + b^2)(a^{-2} + b^{-2}) + 4] \; .$$

Example 3. Show that

$$\int_C (y^2 + z^2) \, dx + (z^2 + x^2) \, dy + (x^2 + y^2) \, dz = -2\pi a b^2$$

where the curve C is the part for which $z \geq 0$ of the intersection of the surfaces

$$x^2 + y^2 + z^2 = 2ax, \quad x^2 + y^2 = 2bx, \quad a > b > 0$$

- The curve begins at the origin and runs at first in the positive octant. Curve C is the intersection of the two surfaces,

$$(x - b)^2 + y^2 = b^2$$
$$z^2 = 2(a - b)x$$

So, the parametric equation of C is

$$x = b(1 + \cos \theta) = 2b \cos^2 \frac{\theta}{2}$$

$$y = b \sin \theta = 2b \sin \frac{\theta}{2} \cos \frac{\theta}{2}$$

$$z = \sqrt{2(a - b) b (1 + \cos \theta)}$$

$$= 2 \sqrt{b(a - b)} \cos \frac{\theta}{2}$$

and θ varies from π to $-\pi$.

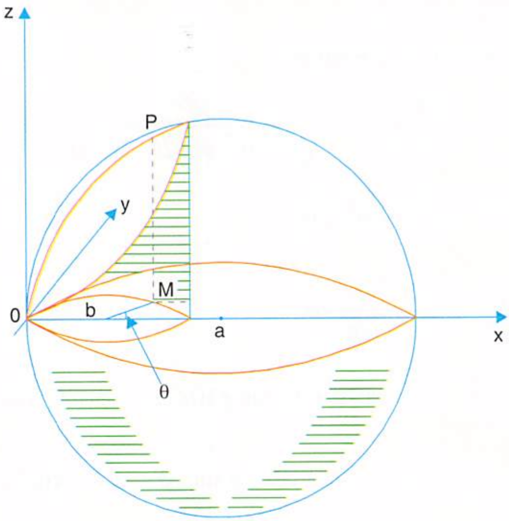


Fig. 1

$$\therefore \qquad y^2 + z^2 = -4b^2 \cos^4 \frac{\theta}{2} + 4ab \cos^2 \frac{\theta}{2}$$

$$z^2 + x^2 = 4ab \cos^2 \frac{\theta}{2} - 4b^2 \cos^2 \frac{\theta}{2} \sin^2 \frac{\theta}{2}$$

$$x^2 + y^2 = 4b^2 \cos^2 \frac{\theta}{2}$$

$$\int_C (y^2 + z^2)\, dx + (z^2 + x^2)\, dy + (x^2 + y^2)\, dz$$

$$= +4b^2 \int_{-\pi}^{\pi} \left(b \cos^4 \frac{\theta}{2} - a \cos^2 \frac{\theta}{2} \right) \sin \theta \, d\theta$$

$$- 4b^2 \int_{-\pi}^{\pi} \left(a \cos^2 \frac{\theta}{2} - b \cos^2 \frac{\theta}{2} \sin^2 \frac{\theta}{2} \right) \cos \theta \, d\theta$$

$$+ 4b^2 \sqrt{b(a-b)} \int_{-\pi}^{\pi} \cos^2 \frac{\theta}{2} \sin \frac{\theta}{2} \, d\theta .$$

As the first and the third integral on the right vanish,

$$= -8b^2 \int_0^{\pi} \left(a \cos^2 \frac{\theta}{2} - b \cos^2 \frac{\theta}{2} \sin^2 \frac{\theta}{2} \right) \cos \theta \, d\theta$$

$$= -8b^2 \int_0^{\pi} a \cos^2 \frac{\theta}{2} \left(2 \cos^2 \frac{\theta}{2} - 1 \right) d\theta$$

$$= -16b^2 \int_0^{\pi/2} a \cos^2 t (2 \cos^2 t - 1)\, dt, \quad t = \frac{\theta}{2} = -2\pi ab^2 .$$

EXERCISE

1. Evaluate the integrals:

   (*i*) $\int \frac{ds}{x-y}$, along the line $2y = x - 4$ between the points $(0, -2)$ and $(4, 0)$.

   (*ii*) $\int y\, ds$, along the arc of the parabola $y^2 = 2px$ cut off by $x^2 = 2py$.

   (*iii*) $\int xy\, ds$, along the ellipse $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$ lying in the first quadrant.

   (*iv*) $\int x\sqrt{x^2 - y^2}\, ds$, along the half-lemniscate $(x^2 + y^2)^2 = a^2(x^2 - y^2)$, $(x \geq 0)$.

   [*Hint*: $x = a\cos^{1/2} 2\theta \cos\theta$, $y = a\cos^{1/2} 2\theta \sin\theta$, $-\pi/4 \leq \theta \leq \pi/4$.]

2. Show that

$$\int_\Gamma xyz\, ds = \frac{\sqrt{3}}{32} R^4$$

   where Γ is the quarter circle of the circle $x^2 + y^2 + z^2 = R^2$, $x^2 + y^2 = R^2/4$ lying in the first octant.

3. Show that

$$\int_r (x + y)\, ds = \sqrt{2}a^2$$

   where Γ is the quarter $x^2 + y^2 + z^2 = a^2$, $y = x$ lying in the first octant.

4. Evaluate the following integrals along segment of straight lines joining the given points.

   (*i*) $\int x\, dx + y\, dy + (x + y - 1)\, dz$, $(1, 1, 1)$ and $(2, 3, 4)$.

   (*ii*) $\int \frac{x\, dx + y\, dy + z\, dz}{\sqrt{(x^2 + y^2 + z^2 - x - y - 2z)}}$, $(1, 1, 1)$ and $(4, 4, 4)$.

5. Find the line integral

$$\int_C (y + z)\, dx + (z + x)\, dy + (x + y)\, dz$$

   where C is the circle $x^2 + y^2 + z^2 = a^2$, $x + y + z = 0$.

6. Evaluate:

$$\int_C x^2 y^3\, dx + dy + z\, dz,$$

   where C is the circle $x^2 + y^2 = R^2$, $z = 0$.

7. Show that

$$\int_\Gamma yz\, dx + zx\, dy + xy\, dz = 0$$

   where Γ is the arc of the curve $x = b\cot t$, $y = b\sin t$, $z = at/2\pi$, from the point it intersects $z = 0$ to the point it intersects $z = a$.

8. Show that

$$\int_C y^2\, dx + z^2\, dy + x^2\, dz = -\frac{\pi a^3}{4}$$

   where C is the curve of intersection of the sphere $x^2 + y^2 + z^2 = a^2$ and the cylinder $x^2 + y^2 = ax$, $(a > 0, z \geq 0)$, integrated anticlockwise when viewed from the origin.

CHAPTER 18

ANSWERS

1. (i) $\sqrt{5}\log 2$ (ii) $(5\sqrt{5}-1)\dfrac{p^2}{3}$ (iii) $\dfrac{ab(a^2+ab+b^2)}{3(a+b)}$ (iv) $\dfrac{2\sqrt{2a^3}}{3}$

4. (i) 13 (ii) $3\sqrt{3}$ 5. 0 6. $\dfrac{-\pi R^6}{8}$

3. SURFACES

A curve in R^3 is a vector valued function whose domain is a subset of R and range a subset of R^3. This idea is extended to define a surface in R^3.

Definition. *A surface in R^3 is a vector-valued function with domain a subset of R^2 and range a subset of R^3.*

Very often we do not make a distinction between a surface and its range set, and take the range of the surface as the surface itself. *Thus if X, Y, Z are three real-valued functions defined on a domain $E \subset R^2$, where*

$$x = X(u, v),\ y = Y(u, v),\ z = Z(u, v),\ (u, v) \in E,$$

then the set

$$\{(X(u, v), Y(u, v), Z(u, v) : (u, v) \in E\} \qquad ...(1)$$

is a (parametric representation of) surface in R^3.

Thus while a curve requires one parameter, the surface requires two parameters for its representation.

A surface of the form

$$z = \Psi(x, y) \quad x \in R,\ y \in R \qquad ...(2)$$

which is met in not more than one point by any line parallel in the axis of z, is known as quadratic or regular with respect to z-axis. Such a surface is projectable in a one-to-one manner on the xy-plane.

$x = \theta(y, z)$ and $y = \phi(z, x)$ are surfaces quadratic (regular) with respect to x-axis and y-axis, respectively.

A surface which can be divided by a finite number of smooth curves into a finite number of portions each of which is quadratic (regular) with respect to the axes, is called a *piecewise quadratic surface.*

A surface is said to be smooth if X, Y, Z possess continuous first order partial derivatives at each point of E and

$$\frac{\partial(y,z)}{\partial(u,v)}, \frac{\partial(z,x)}{\partial(u,v)}, \frac{\partial(x,y)}{\partial(u,v)}$$

do not all vanish simultaneously at any point.

3.1 The Surface Area

Let it be required to compute the area S of a surface bounded by a curve C. The surface being defined by the equation

$$z = \Psi(x, y)$$

where the function Ψ is continuous and possesses continuous partial derivatives. Denote the projection of C on the xy-plane bounded by Γ, and let D be the domain on the xy-plane bounded by Γ. Let σ denote the area of D.

Let an arbitrary partition P of D give rise to sub-domains of areas $\Delta\sigma_1, \Delta\sigma_2, \ldots, \Delta\sigma_n$. In each sub-domain take a point $P_i(\xi_i, \eta_i)$. To the point P_i, there will correspond, on the surface, a point $Q_i(\xi_i, \eta_i, \Psi(\xi_i, \eta_i))$. Through Q_i draw a tangent plane to the surface, and on this plane pick out a sub-domain ΔS_i which projects onto $\Delta\sigma_i$ on the xy-plane.

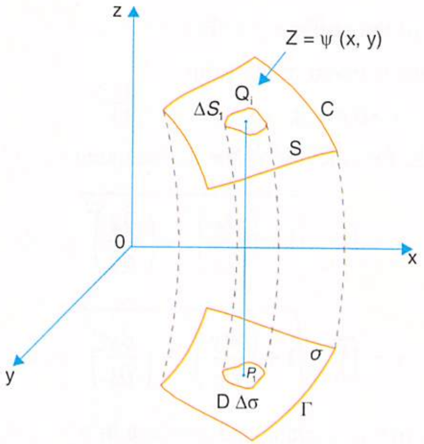


Fig. 2

Consider the sum $\sum\limits_{i=1}^{n} \Delta S_i$.

If the limit of this sum exists, when the norm $\mu(P)$ of the partition (the greatest of the diameters of the sum-domains) approaches zero, the surface is said to be *Squarable* and its area S is given by

$$S = \lim_{\mu(P)\to 0} \sum_{i=1}^{n} \Delta S_i$$

Let us now calculate the value of S.

If γ_i denote the angle between the tangent plane (at Q_i) and the xy-plane, we know from analytical geometry that

$$\Delta\sigma_i = \Delta S_i \cos\gamma_i$$

and

$$\cos\gamma_i = \frac{1}{\sqrt{1 + \Psi_x^2(\xi_i, \eta_i) + \Psi_y^2(\xi_i, \eta_i)}}$$

$$\therefore \qquad \Delta S_i = \sqrt{1 + \Psi_x^2(\xi_i, \eta_i) + \Psi_y^2(\xi_i, \eta_i)}\ \Delta\sigma_i$$

CHAPTER 18

$$\therefore \qquad S = \lim_{\mu(P)\to 0} \sum_i \Delta S_i$$

$$= \lim_{\mu(P)\to 0} \sum_i \sqrt{1 + \Psi_x^2(\xi_i, \eta_i) + \Psi_y^2(\xi_i, \eta_i)}\ \Delta\sigma_i$$

$$= \iint_D \sqrt{1 + \left(\frac{\partial z}{\partial x}\right)^2 + \left(\frac{\partial z}{\partial y}\right)^2}\ dx\, dy \qquad ...(1)$$

a formula to compute the area of the surface $z = \Psi(x, y)$.

If the equation of the surface is given in the form

$$x = \theta(y, z), \text{ or } y = \phi(z, x)$$

then the corresponding formulas for calculating the surface area are of the form

$$S = \iint_{D_1} \sqrt{1 + \left(\frac{\partial x}{\partial y}\right)^2 + \left(\frac{\partial x}{\partial z}\right)^2}\ dy\, dz \qquad ...(2)$$

$$S = \iint_{D_2} \sqrt{1 + \left(\frac{\partial y}{\partial z}\right)^2 + \left(\frac{\partial y}{\partial x}\right)^2}\ dz\, dx \qquad ...(3)$$

where D_1, D_2 are the domains in the *yz*-plane and *zx*-plane in which the given surface is projected.

3.2 The formulas obtained in the above section enable us to find the areas of smooth surfaces (of the form $z = \Psi(x, y)$) which are projectable in a one-to-one manner on the coordinate planes.

However, they help us to deduce a formula to find the area of a smooth surface represented parametrically which is not necessarily projectable in a one-to-one manner on one of the coordinate planes.

To find the area of a smooth surface represented parametrically as

$$x = X(u, v),\ \ y = Y(u, v),\ \ z = Z(u, v),\ \ (u, v) \in E$$

we have only to change the variables in formula (1) § 3.1, where the equation of surface is $z = \Psi(x, y)$.

Put $\qquad x = X(u, v),\ y = Y(u, v)$ so that

$$z = \Psi[X(u, v), Y(u, v)] = Z(u, v)$$

Also (by § 8 Ch. 11),

$$\frac{\partial z}{\partial x} = -\frac{\partial(Y, Z)}{\partial(u, v)} \Bigg/ \frac{\partial(X, Y)}{\partial(u, v)};\quad \frac{\partial z}{\partial y} = -\frac{\partial(Z, X)}{\partial(u, v)} \Bigg/ \frac{\partial(X, Y)}{\partial(u, v)}$$

and then

$$\sqrt{1 + \left(\frac{\partial z}{\partial x}\right)^2 + \left(\frac{\partial z}{\partial y}\right)^2} = \sqrt{\left[\frac{\partial(X, Y)}{\partial(u, v)}\right]^2 + \left[\frac{\partial(Y, Z)}{\partial(u, v)}\right]^2 + \left[\frac{\partial(Z, X)}{\partial(u, v)}\right]^2} \Bigg/ \frac{\partial(X, Y)}{\partial(u, v)}$$

and the Jacobian of transformation $J = \dfrac{\partial(X, Y)}{\partial(u, v)}$.

Thus formula (1) yields

$$S = \iint_D \sqrt{\left[\frac{\partial(X, Y)}{\partial(u, v)}\right]^2 + \left[\frac{\partial(Y, Z)}{\partial(u, v)}\right]^2 + \left[\frac{\partial(Z, X)}{\partial(u, v)}\right]^2}\, du\, dv \qquad \ldots(4)$$

which is the required formula.

It can be easily verified that

$$\left.\begin{aligned}
&\left[\frac{\partial(X, Y)}{\partial(u, v)}\right]^2 + \left[\frac{\partial(Y, Z)}{\partial(u, v)}\right]^2 + \left[\frac{\partial(Z, X)}{\partial(u, v)}\right]^2 = AC - B^2 \\
\text{where}\quad &\left(\frac{\partial X}{\partial u}\right)^2 + \left(\frac{\partial Y}{\partial u}\right)^2 + \left(\frac{\partial Z}{\partial u}\right)^2 = A \\
&\frac{\partial X}{\partial u}\frac{\partial X}{\partial v} + \frac{\partial Y}{\partial u}\frac{\partial Y}{\partial v} + \frac{\partial Z}{\partial u}\frac{\partial Z}{\partial v} = B \\
&\left(\frac{\partial X}{\partial v}\right)^2 + \left(\frac{\partial Y}{\partial v}\right)^2 + \left(\frac{\partial Z}{\partial v}\right)^2 = C
\end{aligned}\right\} \qquad \ldots(5)$$

Thus (4) may be expressed as

$$S = \iint_D \sqrt{AC - B^2}\, du\, dv\,. \qquad \ldots(6)$$

Note: The following results of differential geometry are very useful.

1. (*a*) For the surface $x = X(u, v)$, $y = Y(u, v)$, $z = Z(u, v)$ the direction cosines of the normal to the surface at any point (x, y, z) are

$$\frac{\partial(y, z)}{\partial(u, v)}\Big/\sqrt{K},\ \frac{\partial(z, x)}{\partial(u, v)}\Big/\sqrt{K},\ \frac{\partial(x, y)}{\partial(u, v)}\Big/\sqrt{K}$$

where
$$\begin{aligned}K &= \left[\frac{\partial(y, z)}{\partial(u, v)}\right]^2 + \left[\frac{\partial(z, x)}{\partial(u, v)}\right]^2 + \left[\frac{\partial(x, y)}{\partial(u, v)}\right]^2 \\ &= AC - B^2 \quad \text{[ref. (5) § 3.2]}\end{aligned}$$

(*b*) The elementary surface area dS is given by

$$dS = \sqrt{K}\, du\, dv.$$

2. (*a*) For the surface $z = \Psi(x, y)$, the direction cosines of the normal at any point (x, y, z) of the surface area are

$$-\frac{\partial z}{\partial x}\Big/\sqrt{K},\ -\frac{\partial z}{\partial y}\Big/\sqrt{K},\ 1/\sqrt{K}\,.$$

where $$K = 1 + \left(\frac{\partial z}{\partial x}\right)^2 + \left(\frac{\partial z}{\partial y}\right)^2$$

(*b*) The elementary surface area dS is given by

$$dS = \sqrt{K}\ dx\,dy = \sqrt{1 + \left(\frac{\partial z}{\partial x}\right)^2 + \left(\frac{\partial z}{\partial y}\right)^2}\ dx\,dy.$$

Example 4. Compute the surface area S of the sphere

$$x^2 + y^2 + z^2 = a^2$$

■ The surface area of the sphere is twice the surface area of the upper half-sphere, $z = \sqrt{a^2 - x^2 - y^2}$. Now

$$\frac{\partial z}{\partial x} = -\frac{x}{\sqrt{a^2 - x^2 - y^2}},\quad \frac{\partial z}{\partial y} = -\frac{y}{\sqrt{a^2 - x^2 - y^2}}$$

so that

$$\sqrt{1 + \left(\frac{\partial z}{\partial x}\right)^2 + \left(\frac{\partial z}{\partial y}\right)^2} = \frac{a}{\sqrt{a^2 - x^2 - y^2}}$$

The domain of integration is the circle $x^2 + y^2 = a^2$ on the xy-plane.

Thus by formula (1), we have

$$\tfrac{1}{2}S = \int_{-a}^{a} dx \int_{-\sqrt{(a^2-x^2)}}^{\sqrt{(a^2-x^2)}} \frac{a}{\sqrt{a^2 - x^2 - y^2}}\,dy$$

On passing to polars, we have

$$S = 2\int_0^{2\pi}\left[\int_0^a \frac{a}{\sqrt{a^2 - r^2}}\,r\,dr\right]d\theta = 4\pi a^2.$$

Example 5. Find the area of that part of the surface of the cylinder $x^2 + y^2 = a^2$ which is cut out by the cylinder $x^2 + z^2 = a^2$.

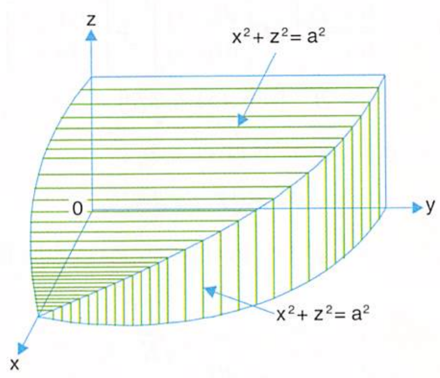


Fig. 3

The figure shows 1/8th of the desired surface.

The equation of the surface has the form

$$y = \sqrt{a^2 - x^2}$$

so that

$$\frac{\partial y}{\partial x} = -\frac{x}{\sqrt{a^2 - x^2}}, \quad \frac{\partial y}{\partial z} = 0$$

$$\sqrt{1 + \left(\frac{\partial y}{\partial x}\right)^2 + \left(\frac{\partial y}{\partial z}\right)^2} = \frac{a}{\sqrt{a^2 - x^2}}$$

The domain of integration is a quarter circle $x^2 + z^2 = a^2, x \geq 0, z \geq 0,$ on the xz-plane. Thus by formula (3), we have

$$\frac{1}{8} S = \int_0^a dx \int_0^{\sqrt{a^2 - x^2}} \frac{a}{\sqrt{a^2 - x^2}} dz = a^2$$

$\therefore$ $$S = 8a^2.$$

Example 6. The x and y coordinates of a point on the paraboloid $2z = x^2/a + y^2/b$ are expressed in the form

$$x = a \tan \theta \cos \phi, \; y = b \tan \theta \sin \phi$$

where θ is the angle of inclination of the normal at any point to the axis of z. Show that the area of the cap of the surface cut off by the curve $\theta = \lambda$ is $2\pi ab(\sec^3 \lambda - 1)/3$.

■ Here

$$1 + \left(\frac{\partial z}{\partial x}\right)^2 + \left(\frac{\partial z}{\partial y}\right)^2 = 1 + \frac{x^2}{a^2} + \frac{y^2}{b^2}$$

so that by formula (1), surface area S is given by

$$S = \iint_D \sqrt{1 + \frac{x^2}{a^2} + \frac{y^2}{b^2}} \, dx \, dy$$

Let us put $x = a \tan \theta \cos \phi, y = b \tan \theta \sin \phi$.

$\therefore$ $$\sqrt{1 + \frac{x^2}{a^2} + \frac{y^2}{b^2}} = \sec \theta$$

and

$$J = \frac{\partial(x, y)}{\partial(\theta, \phi)} = ab \sec^2 \theta \tan \theta$$

$\therefore$ $$S = \int_0^{2\pi} d\phi \int_0^{\lambda} \sec \theta \, ab \sec^2 \theta \tan \theta \, d\theta$$

$$= 2\pi \int_0^{\lambda} ab \sec^3 \theta \tan \theta \, d\theta = \frac{2\pi ab}{3} (\sec^3 \lambda - 1).$$

Example 7. A surface is given by the equations

$$x = c \sin u, \quad y = c \cos v, \quad z = c(\cos u + \cos v)$$

prove that its area bounded by

$$u = 0, \quad u = \pi/2; \quad v = 0, \quad v = \pi/2$$

is

$$\frac{\pi c^2}{2} \left[1 - \sum_{n=1}^{\infty} (-1)^n \frac{p_{2n}^2}{2n - 1} \right]$$

where
$$p_{2n} = \frac{(2n-1)(2n-3)\ldots 3.1}{2n(2n-2)\ldots 4.2}$$

■ Here

$$A = c^2 \cos^2 u + c^2 \sin^2 u = c^2$$
$$B = c^2 \sin u \sin v$$
$$C = c^2 \sin^2 v + c^2 \sin^2 v = 2c^2 \sin^2 v$$

$$\therefore \quad S = \iint_D c^2 \sqrt{2 \sin^2 v - \sin^2 u \sin^2 v} \, du \, dv$$

$$= c^2 \int_0^{\pi/2} \sin v \, dv \int_0^{\pi/2} \sqrt{1 + \cos^2 u} \, du = c^2 \int_0^{\pi/2} (1 + \cos^2 u)^{1/2} \, du$$

$$= c^2 \int_0^{\pi/2} \left[1 + \frac{1}{2} \cos^2 u + \frac{\frac{1}{2}\left(\frac{1}{2} - 1\right)}{2!} (\cos^2 u)^2 + \ldots + \frac{\frac{1}{2}\left(\frac{1}{2} - 1\right)\ldots\left(\frac{1}{2} - n + 1\right)}{n!} (\cos^2 u)^2 + \ldots \right] du$$

$$= c^2 \int_0^{\pi/2} \left[1 + \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{2n-1} \frac{(2n-1)(2n-3)\ldots 3.1}{2n(2n-2)\ldots 4.2} \cos^{2n} u \right] du$$

$$= \frac{\pi}{2} c^2 \left[1 - \sum_{n=1}^{\infty} \frac{(-1)^n p_{2n}^2}{2n-1} \right].$$

EXERCISE

1. Show that the area of the surface of the sphere $x^2 + y^2 + z^2 = a^2$ cut off by $x^2 + y^2 = ax$ is $2(\pi - 2)a^2$.
2. Find the area of the part of the spherical surface

$$x^2 + y^2 + z^2 = 4a^2$$

enclosed by the cylinder

$$(x^2 + y^2)^2 = 2a^2(2x^2 + y^2).$$

3. Find the areas of the indicated parts of the given surfaces.
   (i) The part of $z^2 = x^2 + y^2$ cut off by the cylinder $z^2 = 2py$.
   [*Hint*: Take projection on xy-plane.]
   (ii) The part of $y^2 + z^2 = x^2$ inside the cylinder $x^2 + y^2 = a^2$.
   [*Hint*: Projection on yz-plane.]
   (iii) The part of $y^2 + z^2 = x^2$ cut off by the cylinder $x^2 - y^2 = a^2$ and the planes $y = b,\ y = -b$.
   (iv) The part of $z^2 = 4x$ cut off by the cylinder $y^2 = 4x$ and the plane $x = 1$.
   (v) The part of $z = xy$ cut off by the cylinder $x^2 + y^2 = a^2$.
   (vi) The part of $2z = x^2 + y^2$ cut off by the cylinder $x^2 + y^2 = 1$.
   (vii) The part of the cone $z^2 = x^2 + y^2$ inside the cylinder $x^2 + y^2 = 2x$.
   (viii) The part of $x^2 = y^2 + z^2$ between the cylinder $y^2 = z$ and the plane $y = z - 2$.
   (ix) The part of the cone $x^2 = y^2 + z^2$ inside the sphere $x^2 + y^2 + z^2 = 2z$.
   (x) The part of $y^2 + z^2 = 2z$ cut off by the cone $x^2 = y^2 + z^2$.
4. Find the area of the surface of the cylinder $x^2 + y^2 = 4a^2$ above the xy-plane and bounded by the planes $y = 0$ $z = a$ and $y = z$.
5. Find the surface area of the part of the sphere $x^2 + y^2 + z^2 = a^2$ above $z = 0$ plane, cut off by the vertical cylinder erected on one loop of the curve whose equation in polars is $r = a \cos 2\theta$.
6. Show that the area of the surface of the paraboloid $\dfrac{x^2}{a} + \dfrac{y^2}{b} = 2z$ inside the cylinder $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = k$ is $\dfrac{2}{3}\pi\{(1+k)^{3/2} - 1\}\, ab$.
7. Show that the surface area of the sphere $x^2 + y^2 + z^2 = 1$ that lies inside the cylinder $2x^2(x^2 + y^2) = 3(x^2 - y^2)$ is $2\pi - 4\sqrt{2}\{\sqrt{3}\log(\sqrt{3} + \sqrt{2}) - 2\log(1 + \sqrt{2})\}$.
8. Calculate the area of the spherical surface given by
$$x = a\cos\theta\cos\phi,\quad y = a\cos\theta\sin\phi,\quad z = a\sin\theta$$
   where $0 < \phi < 2\pi,\ -\dfrac{\pi}{2} < \theta < \dfrac{\pi}{2}$
9. Find the surface area of the torus
$$x = (r - \cos v)\cos u,\quad y = (r - \cos v)\sin u,\quad z = \sin v,$$
   where $-\pi \le u, v \le \pi,\quad r > 1.$
10. Compute the area of the part of the earth's surface (considering it as a sphere of radius R, km) contained between the meridians $\phi = 30^\circ,\ \phi = 60^\circ$ and parallels $\theta = 45^\circ,\ \theta = 60^\circ$.

ANSWERS

2. $16(\pi - \sqrt{2})a^2$
3. (i) $2\sqrt{2}\pi p^2$ (ii) $2\pi a^2$
   (iii) $8\sqrt{2}ab$ (iv) $16(\sqrt{8} - 1)/3$
   (v) $\dfrac{2\pi}{3}\{(1 + a^2)^{3/2} - 1\}$ (vi) $2\pi(\sqrt{8} - 1)/3$

(vii) $2\pi\sqrt{2}$ (viii) $9\sqrt{2}$

(ix) $\pi\sqrt{2}$ (x) 16

4. $\frac{a^2}{3}(\pi + 6\sqrt{3} - 12)$

5. $\frac{1}{2}(\pi - 2)a^2$

8. $4\pi a^2$

9. $4\pi^2 r$

10. $\frac{\pi R^2}{12}(\sqrt{3} - \sqrt{2})$

4. SURFACE INTEGRALS

In many physical problems we encounter functions defined on various surfaces, for example, density of a charge distribution over the surface of a conductor, intensity of illumination of a surface, velocity of the particles of a fluid passing through a surface, and the like. This section is devoted to studying integrals of functions defined on surfaces, the so-called surface integrals. The theory is in many respects analogous to the theory of line integrals presented in earlier sections.

4.1 Surface Integrals of the First Type (*Definition*)

Let S be a smooth (or piecewise smooth) surface bounded by a smooth (or piecewise smooth) contour C. Let a bounded function f be defined at all points (x, y, z) of the surface.

Let the surface be partitioned into sub-surfaces of areas $\Delta S_1, \Delta S_2, \ldots, \Delta S_n$ by means of smooth (or piecewise smooth) curves. Let $P(\xi_i, \eta_i, \zeta_i)$ be a point of ΔS_i. For the sum

$$\sum_i f(\xi_i, \eta_i, \zeta_i)\, \Delta S_i . \qquad \ldots(1)$$

If the sum tends to a finite limit as the norm of the partition (the maximal of the diameters of the sub-areas ΔS_i) tends to zero, and for all positions of (ξ_i, η_i, ζ_i) in ΔS_i the limit is called the *surface integral of the first type* of the function f over the surface S and is denoted by the symbol

$$\iint_S f(x, y, z)\, dS \qquad \ldots(2)$$

Note: Let a plane surface D bounded by contour Γ be the projection of the surface S on the xy-plane and $\Delta\sigma_i$ the corresponding sub-areas of D.

Then

$$\Delta\sigma_i = \Delta S_i \cos\gamma_i$$

where γ_i is the angle of inclination of the normal at (ξ_i, η_i, ζ_i) to the surface ΔS_i with the z-axis. Sum (1) then becomes

$$\sum_i f(\xi_i, \eta_i, \zeta_i)\, \Delta S_i = \sum_i f(\xi_i, \eta_i, \zeta_i) \frac{\Delta\sigma_i}{\cos\gamma_i}$$

which, in the limit, yields

$$\iint_S f(x, y, z)\, dS = \iint_D f(x, y, z)\, \frac{dx\, dy}{\cos \gamma} \qquad ...(3)$$

where z is on the right hand side, is expressed in terms of x, y with the help of the equation of the surface.

Relation (3) expresses a surface integral in terms of a double integral.

4.2 Reducing a Surface Integral of First Type to a Double Integral

(*i*) If the surface is represented by $z = \Psi(x, y)$, we know

$$\Delta S_i = \sqrt{1 + \Psi_x^2 + \Psi_y^2}\, \Delta x_i\, \Delta y_i$$

so that in the limit, (1) takes the form

$$\iint_S f(x, y, z)\, dS = \iint_D f(x, y, z) \sqrt{1 + \left(\frac{\partial z}{\partial x}\right) + \left(\frac{\partial z}{\partial y}\right)^2}\, dx\, dy \qquad ...(4)$$

where z, in the double integral, is expressed in terms of x, y and D is the projection of S on the xy-plane.

If the surface is represented by equation of the form

$$x = \theta(y, z) \quad \text{or} \quad y = \phi(z, x)$$

then interchanging the roles of the variables x, y, z, the corresponding formulas are of the form

$$\iint_S f(x, y, z)\, dS = \iint_{D_1} f(\theta(y, z), y, z) \sqrt{1 + \left(\frac{\partial x}{\partial y}\right)^2 + \left(\frac{\partial x}{\partial z}\right)^2}\, dy\, dz \qquad ...(5)$$

$$\iint_S f(x, y, z)\, dS = \iint_{D_2} f(x, \phi(z, x), z) \sqrt{1 + \left(\frac{\partial y}{\partial z}\right)^2 + \left(\frac{\partial y}{\partial x}\right)^2}\, dz\, dx \qquad ...(6)$$

where D_1 and D_2 are respectively the projection of S on the yz-plane and zx-plane.

Note: If the surface S is composed of several parts, each of which can be represented by an equation of the form

$$x = \theta(y, z), \quad y = \phi(z, x) \quad \text{or} \quad z = \Psi(x, y)$$

then, since by definition, the surface integral over S is equal to the sum of the integrals over the parts of S, applying formulas (4), (5) or (6) to these integrals separately, we reduce the integral over S to the sum of the double integrals.

(*ii*) In case a smooth surface S is represented by a parametric equation of the form

$$x = X(u, v), \quad y = Y(u, v), \quad z = Z(u, v)$$

then by applying arguments essentially the same as above, the surface integral of a bounded function f over S can be expressed in terms of a double integral, by the relation

$$\left.\begin{aligned}\iint_S f(x, y, z)\,dS &= \iint_D f(X, Y, Z)\sqrt{K}\,du\,dv \\ \text{or}\qquad &= \iint_D f(X, Y, Z)\sqrt{AC - B^2}\,du\,dv\end{aligned}\right\} \quad ...(7)$$

where the domain D in uv-plane corresponds to S, and

$$K = \left[\frac{\partial(x, y)}{\partial(u, v)}\right]^2 + \left[\frac{\partial(y, z)}{\partial(u, v)}\right]^2 + \left[\frac{\partial(z, x)}{\partial(u, v)}\right]^2$$

$$A = \left(\frac{\partial x}{\partial u}\right)^2 + \left(\frac{\partial y}{\partial u}\right)^2 + \left(\frac{\partial z}{\partial u}\right)^2$$

$$B = \frac{\partial x}{\partial u}\frac{\partial x}{\partial v} + \frac{\partial y}{\partial u}\frac{\partial y}{\partial v} + \frac{\partial z}{\partial u}\frac{\partial z}{\partial v}$$

$$C = \left(\frac{\partial x}{\partial v}\right)^2 + \left(\frac{\partial y}{\partial v}\right)^2 + \left(\frac{\partial z}{\partial v}\right)^2.$$

Note: We have considered surface integrals of scalar function f. This notion can be easily generalized to a vector function $\mathbf{F}$ defined on S. Let

$$\mathbf{F}(x, y, z) = \mathbf{i}\mathbf{P}(x, y, z) + \mathbf{j}\mathbf{Q}(x, y, z) + \mathbf{k}\mathbf{R}(x, y, z)$$

where P, Q, R are scalar functions. We define the surface integral of $\mathbf{F}$ over S by the relation

$$\iint_S \mathbf{F}\,dS = \mathbf{i}\iint_S P\,dS + \mathbf{j}\iint_S Q\,dS + \mathbf{k}\iint_S R\,dS \quad ...(8)$$

where the surface integrals of scalar functions on the right side can be expressed as double integrals, and call it the surface integral of the first type of the vector function $\mathbf{F}$ over the surface S.

We now proceed to discuss the theory of *Surface integrals of the second type*. To understand the theory, let us first discuss the question of choosing a side of a surface, which is analogous to the problem of introducing an orientation of a curve.

4.3 Oriented Surfaces, Positive and Negative Sides

A surface is said to be two sided (*bilateral*) if it is possible to distinguish one of its sides from the other. We assume that a surface has two distinct sides in such a way that it is impossible to pass from one side to the other along a continuous path which lies on the surface and which does not cross one of the bounding curves. However, all surfaces are not two-sided. The simplest example of a unilateral (one-sided) surface is provided by the well known *Möbius strip*, shown in the figure (*c*) which may be obtained by taking a rectangular strip *ABCD* of paper and pasting it on two sides *BC* and *AD* after giving a half twist, *i.e.*, in such a way that the point *A* coincides with the point *C*, and *B* with *D*.

The concept of a side of a surface is closely related to the orientation of its bounding curve. A side of a surface (a region) is said to be positive or positively oriented if the orientation of its boundary is positive.

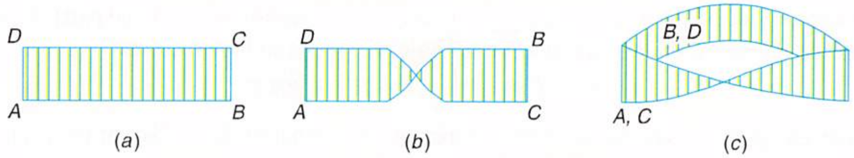


Fig. 4

Let S be a two-sided smooth surface and C its bounding curve. Take a small area σ bounded by a curve Γ on the surface. Take a point P in σ and the normal PNP' to the surface at P. The half lines PN and PN', which do not pierce the surface, are drawn in opposite directions from the surface; one of these directions, say that of PN, is chosen as the positive direction of the normal at P, and PN may be called the positive normal, PN' the negative normal. That side of the surface σ which faces the positive direction of the normal at P will be called the positive (or upper) side of the surface; the other side will be the negative (or lower) side.

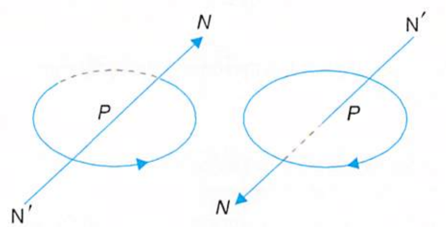


Fig. 5

The concept of a side of a surface is closely related to the orientation of its contour. A side of a surface is said to be positive or positively oriented if the orientation of its boundary is positive.

To correlate these two ideas, we introduce for each contour entering into the boundary its orientation according to the following rule:

'A direction in which the contour C is described is considered to be positive if the surface S is always kept on the left of a person who is on the surface and walking round the contour in this direction, so that the normal PN goes from his feet to his head. The opposite side is referred to as the negative one.'

Thus the side of the surface with this orientation of the boundary is called the positive side and the other the negative side. The two sides of the surface are then called *orientable* and the process of choosing a certain side of a two-sided surface is referred to as the *orientation* of the surface. The one-sided surfaces are thus non-orientable.

Plane surfaces are oriented in the same way. Accordingly the positive side of xy-plane is that which faces the positive direction of z-axis.

In what follows, it will always be assumed that the positive side of a surface projects on the positive side of the coordinate planes.

4.4 Surface Integral of the Second Type: Flux Across a Surface

Definition. Let S be a smooth (or piecewise smooth) two-sided (oriented) surface, bounded by a smooth (or piecewise smooth) contour C. Let a bounded vector-valued function $F = (f, g, h)$ be defined for all

points of a certain side, say positive side of the surface. The projection (or resolved part) F_n of F in the direction of the normal to the surface at arbitrary point (x, y, z) can be written as

$$F_n = f \cos\alpha + g \cos\beta + h \cos\gamma$$

where $\cos\alpha, \cos\beta, \cos\gamma$ are the direction cosines of the normal to the surface at (x, y, z).

The surface integral

$$\iint_S (f \cos\alpha + g \cos\beta + h \cos\gamma)\, dS \qquad \text{...(1)}$$

is called the *Surface integral of the second type* of the vector valued function $F = (f, g, h)$ over the surface S (or strictly speaking, over the chosen side of the surface) and will be denoted as

$$\iint_S (f\, dy\, dz + g\, dz\, dx + h\, dx\, dy) \qquad \text{...(2)}$$

Thus by definition, we have the relation

$$\iint_S (f \cos\alpha + g \cos\beta + h \cos\gamma)\, dS = \iint_S f\, dy\, dz + g\, dz\, dx + h\, dx\, dy. \qquad \text{...(3)}$$

Note: The surface integral of the second type of F is the same as the surface integral of first type of F_n (the component of F along the normal to the surface).

Remark: We have defined the surface integral of the second type on the basis of the notion of the surface integral of the first type. But it can be defined directly also as follows:

For brevity, let us first consider only one of the components, say h, of the vector valued functions. Let D be the projection of S on the xy-plane. Since S is positively oriented, D should also be oriented positively. Clearly D is a plane surface on the xy-plane. Let S be partitioned into the sub-areas of D. Choose an arbitrary point (ξ_i, η_i, ζ_i) in each sub-area ΔS_i and consider the sum

$$\sum_{i=1}^{n} h(\xi_i, \eta_i, \zeta_i)\, \Delta\sigma_i \qquad \text{...(4)}$$

If the limit of the sum exists, when the norm of the partition tends to zero (or $n \to \infty$) (which always exists for continuous function h and smooth surface S), the limit is equal to the integral

$$\iint_S h(x, y, z)\, dx\, dy \qquad \text{...(5)}$$

We can similarly define by means of corresponding sums of the integrals

$$\iint_S f(x,y,z) dy\, dz \ \ and \ \iint_S g(x,y,z) dz\, dx$$

and consequently, write the sum of these integrals,

$$\iint_S f\, dy\, dz + g\, dz\, dx + h\, dx\, dy \qquad \text{...(6)}$$

called the *surface integral of the second type* of $F = (f, g, h)$ over S.

Notes:

1. The sum (4) can also be written as

$$\sum_{i=1}^{n} h(\xi_i, \eta_i, \zeta_i) \cos\gamma\, \Delta S_i \qquad \text{...(7)}$$

which represents the flux of h across the surface S, and in the limit, represents the surface integral.

$$\iint_S h \cos \gamma \, dS$$

Thus, in the limit, the sum (4) represents the surface integral, $\iint_S h \, dx \, dy$ as well as $\iint_S h \cos \gamma \, dS$. Proceeding, similarly, for the other two functions f and g, we establish the justification of relation (3).

2. The surface integral of the second type, given by (1), is the flux of $F = (f, g, h)$ across the surface S and that is the reason, the surface integral of the second type is also called the flux of a function across a surface.

4.5 Reducing a Surface Integral of the Second Type to a Double Integral

The definition of the surface integral of the second type implies the following results:

(*i*) Let S be a smooth surface determined by an equation

$$Z = \Psi(x, y)$$

and let $h(x, y, z)$ be a bounded function defined on S. Then for the surface integral of the second type taken over the positive side of S we have the relation (definition, relation (4))

$$\iint_S h(x, y, z) \, dx \, dy, = \iint_D h(x, y, \Psi(x, y)) \, dx \, dy \quad \text{...(8)}$$

where D is the projection of S on the $z = 0$ plane, and since S is positively oriented, the sign before the double integral is to be positive if D is positively oriented, otherwise negative.

If the integral is taken over the other side (negative) of S, the sign before the double integral is to be negative if D is positively oriented, otherwise positive.

We, similarly derive the formula

$$\iint_S f(x, y, z) \, dy \, dz = \iint_{D_1} f(\theta(y, z), y, z) \, dy \, dz \quad \text{...(9)}$$

$$\iint_S g(x, y, z) \, dz \, dx = \iint_{D_2} g(x, \phi(z, x), z) \, dz \, dx \quad \text{...(10)}$$

where D_1, D_2 are the projections of S on yz and zx planes respectively. So finally we have

$$\iint_S f \, dy \, dz + g \, dz \, dx + h \, dx \, dy = \iint_{D_1} f(\theta(y, z), y, z) \, dy \, dz$$

$$+ \iint_{D_2} g(x, \phi(z, x), z) \, dz \, dx + \iint_D h(x, y, \Psi(x, y)) \, dx \, dy \quad \text{...(11)}$$

(*ii*) Let S be a smooth surface represented as

$$x = X(u, v), \quad y = Y(u, v), \quad z = Z(u, v); (u, v) \in D$$

where D is the surface in uv-plane corresponding to the surface S in the xy-plane.

Let $F = (f, g, h)$ be a bounded vector valued function defined on S. Then for the surface integral of the second type taken over the positive side of S, we have the relations (definition, relation (4))

$$\iint_S h(x, y, z) \, dx \, dy = \iint_D h(x, y, z) \, dx \, dy = \iint_D h(X, Y, Z) \frac{\partial(X, Y)}{\partial(u, v)} \, du \, dv.$$

$$\iint_S f(x, y, z)\, dy\, dz = \iint_D f(X, Y, Z)\, \frac{\partial(Y, Z)}{\partial(u, v)}\, du\, dv$$

$$\iint_S g(x, y, z)\, dz\, dx = \iint_D g(X, Y, Z)\, \frac{\partial(Z, X)}{\partial(u, v)}\, du\, dv$$

where D is also oriented in the same sense as S.

Thus, finally we have

$$\iint_S f\, dy\, dz + g\, dz\, dx + h\, dx\, dy$$

$$= \iint_D \left[f\, \frac{\partial(Y, Z)}{\partial(u, v)} + g\, \frac{\partial(Z, X)}{\partial(u, v)} + h\, \frac{\partial(X, Y)}{\partial(u, v)} \right] du\, dv \qquad ...(12)$$

Note: Vectorial formation

Let

$$\mathbf{r} = \mathbf{i}x + \mathbf{j}y + \mathbf{k}z$$

be the position vector of any point on the surface S, and

$$\mathbf{F}(x, y, z) = \mathbf{i}P(x, y, z) + \mathbf{j}Q(x, y, z) + \mathbf{k}R(x, y, z)$$

be a vector function, with P, Q, R as its components.

Let $\mathbf{n}$ denote the unit vector along the normal at any point on the side of the surface under consideration, so that

$$\mathbf{n} = \mathbf{i} \cos\alpha + \mathbf{j} \cos\beta + \mathbf{k} \cos\gamma$$

$$\therefore \qquad \iint_S \mathbf{F}\cdot\mathbf{n}\, dS = \iint_S (P\cos\alpha + Q\cos\beta + R\cos\gamma)\, dS \qquad ...(13)$$

Thus the surface integral of the second type of $\mathbf{F}$ (or the surface integral of the first type of $\mathbf{F}\cdot\mathbf{n}$) over S is $\iint_S \mathbf{F}\cdot\mathbf{n}\, dS$.

Ex. Show that relation (12) satisfies (13).

4.6 Properties of Surface Integrals

1. The surface integrals (of the second type) taken over the opposite sides of a surface have different signs, *i.e.*,

$$\iint_S \mathbf{F}\cdot\mathbf{n}\, dS = -\iint_{S'} \mathbf{F}\cdot\mathbf{n}\, dS$$

where S and S' are the two sides of the surface.

2. If a surface S is broken up into m parts $S_1, S_2, \ldots, S_m$, the integral over the whole surface S (say, over its positive side) is equal to the sum of the integrals taken over the corresponding (*i.e.*, positive) sides of the surfaces $S_1, S_2, \ldots, S_m$.

$$\iint_S \mathbf{F}\cdot\mathbf{n}\, dS = \iint_{S_1} \mathbf{F}\cdot\mathbf{n}\, dS + \iint_{S_2} \mathbf{F}\cdot\mathbf{n}\, dS + \ldots + \iint_{S_m} \mathbf{F}\cdot\mathbf{n}\, ds \qquad ...(14)$$

4.7 Relation Between the Two Types of Surface Integrals

Let S be a smooth (or piecewise smooth) oriented surface and C its bounding curve. Then the surface integral of the second type of a function P, taken over a certain side (say, positive) of the surface is

$\iint_S f\, dy\, dz$, which is equivalent to $\iint_S f \cos\alpha\, dS$. The latter being the surface integral of the first type of the function $f\cos\alpha$.

Similarly, the surface integral of the second type of the function $F = (f, g, h)$ is

$$\iint_S (f\, dy\, dz + g\, dz\, dx + h\, dx\, dy) \qquad ...(1)$$

which is equivalent to

$$\iint_S (f \cos\alpha + g \cos\beta + h \cos\gamma)\, dS \qquad ...(2)$$

The latter being a surface integral of the first type of the function $(f \cos\alpha + g\cos\beta + h\cos\gamma)$. Also $(f\cos\alpha + g\cos\beta + h\cos\gamma)$ is the projection or the component of the function F along the normal to the surface.

Thus the surface integral of the second type of a function is same as the surface integral of the first type of the normal component of the function.

Note: In symbolic representation of the surface integral,

$$\iint_S f\, dS,\ \iint_S f \cos\alpha\, dS \text{ or } \iint_S f\, dy\, dz$$

the meanings are well understood and so there is no need to mention the type.

But when symbols are not being employed, mentioning the type (of the integral) has to be made. Thus for $\iint_S f\, dS$, we say 'the surface integral of the first type of f', and for $\iint f \cos\alpha\, dS$, 'the surface integral of the first type of $f \cos\alpha$ or the surface integral of the second type of f'.

CHAPTER 18

Example 8. Find the value of the surface integral $\iint_S \frac{dS}{r}$, where S is the portion of the surface of the hyperbolic paraboloid $z = xy$ cut off by the cylinder $x^2 + y^2 = a^2$, and r is the distance from a point on the surface to the z-axis.

■ For any point (x, y) on the surface, $r = \sqrt{x^2 + y^2}$.

Now

$$\iint_S \frac{dS}{\sqrt{x^2+y^2}} = \iint_D \frac{1}{\sqrt{x^2+y^2}} \sqrt{1 + \left(\frac{\partial z}{\partial x}\right)^2 + \left(\frac{\partial z}{\partial y}\right)^2}\, dx\, dy$$

where D is the projection of E on xy-plane

$$= \iint_D \frac{\sqrt{1+x^2+y^2}}{\sqrt{x^2+y^2}}\, dx\, dy = \int_0^{2\pi} d\theta \int_0^a \frac{\sqrt{1+r^2}}{r}\, r\, dr$$

$$= \pi\left[a\sqrt{a^2+1} + \log\left(a + \sqrt{a^2+1}\right)\right]$$

Example 9. Evaluate the surface integral

$$\iint_S p(x^4 + y^4 + z^4)\, dS$$

where S is the surface $x^2/a^2 + y^2/b^2 + z^2/c^2 = 1$, and p is the perpendicular from the origin to the tangent plane at the point (x, y, z).

■ Symmetry of the integrand and the surface S show that its value is 8 times the value of the integral over the portion of the surface in the first octant.

$$\text{Let} \quad \left.\begin{aligned} x &= a\cos\theta\cos\phi \\ y &= b\cos\theta\sin\phi \\ z &= c\sin\theta \end{aligned}\right\}, \quad \begin{aligned} 0 \le \phi \le \pi/2 \\ 0 \le \theta \le \pi/2 \end{aligned}$$

Then

$$\frac{\partial(x, y)}{\partial(\theta, \phi)} = -\, ab\cos\theta\sin\theta$$

$$\frac{\partial(y, z)}{\partial(\theta, \phi)} = -\, bc\cos^2\theta\cos\phi$$

$$\frac{\partial(z, x)}{\partial(\theta, \phi)} = -\, ac\cos^2\theta\sin\phi$$

so that

$$\left[\frac{\partial(x, y)}{\partial(u, v)}\right]^2 + \left[\frac{\partial(y, z)}{\partial(u, v)}\right]^2 + \left[\frac{\partial(z, x)}{\partial(u, v)}\right]^2$$

$$= a^2b^2c^2\cos^2\theta\left\{\frac{\cos^2\theta\cos^2\phi}{a^2} + \frac{\cos^2\theta\sin^2\phi}{b^2} + \frac{\sin^2\theta}{c^2}\right\}$$

and

$$p = 1\Big/\sqrt{\frac{\cos^2\theta\cos^2\phi}{a^2} + \frac{\cos^2\theta\sin^2\phi}{b^2} + \frac{\sin^2\theta}{c^2}}$$

The surface integral becomes

$$\iint_S p(x^4 + y^4 + z^4)\, dS$$

$$= 8\int_0^{\pi/2} d\theta \int_0^{\pi/2} abc\cos\theta[a^4\cos^4\theta\cos^4\phi + b^4\cos^4\theta\sin^4\phi + c^4\sin^4\theta]\, d\phi$$

$$= 8abc\int_0^{\pi/2}\cos\theta\left[a^4\cos^4\theta\cdot\frac{3\cdot 1}{4\cdot 2}\,\frac{\pi}{2} + b^4\cos^4\theta\cdot\frac{3\cdot 1}{4\cdot 2}\,\frac{\pi}{2} + c^4\sin^4\theta\cdot\frac{\pi}{2}\right] d\theta$$

$$= 8abc\cdot\frac{\pi}{2}\left[\frac{3}{8}(a^4 + b^4)\frac{4\cdot 2}{5\cdot 3} + \frac{c^4}{5}\right] = \frac{4\pi abc(a^4 + b^4 + c^4)}{5}.$$

Example 10. Evaluate

$$I = \iint_S (x\, dy\, dz + dz\, dx + xz^2\, dx\, dy),$$

where S is the outer side of the part of the sphere $x^2 + y^2 + z^2 = 1$ in the first octant.

- Let us denote the projections of the surface S on the coordinate planes yz, zx and xy by D_1, D_2, D_3 respectively. These are quarter circles of radius 1. Then we have

$$I_1 = \iint_S x\, dy\, dz = \iint_{D_1} \sqrt{1 - y^2 - z^2}\, dy\, dz$$

$$I_2 = \iint_S dz\, dx = \iint_{D_2} dz\, dx$$

$$I_3 = \iint_S xz^2\, dx\, dy = \int_{D_3} x(1 - x^2 - y^2)\, dx\, dy.$$

The second integral I_2 is simply the area of the domain D_2, (quarter circle $x^2 + z^2 = 1$), *i.e.*, equal to $\pi/4$.

$$I_1 = \int_0^{\pi/2} d\theta \int_0^1 \sqrt{1 - r^2}\, r\, dr = \frac{\pi}{6}$$

$$I_3 = \int_0^{\pi/2} \cos\theta\, d\theta \int_0^1 r(1 - r^2)\, r\, dr = \frac{2}{15}$$

Thus,

$$I = I_1 + I_2 + I_3 = \frac{\pi}{6} + \frac{\pi}{4} + \frac{2}{15} = \frac{5\pi}{12} + \frac{2}{15}.$$

Example 11. Evaluate the surface integral $\iint z \cos\gamma\, dS$, over the outer side of the sphere $x^2 + y^2 + z^2 = 1$, where γ is the inclination of the normal at any point of the sphere with the z-axis.

- Here the z-coordinate can be expressed as a single-valued function of x and y for the whole surface. Let us break it into two parts—the upper hemisphere S_1, above the xy-plane, and the lower hemisphere S_2, below it.

  Accordingly their equations are

$$z = \sqrt{1 - x^2 - y^2},\quad z = -\sqrt{1 - x^2 - y^2}$$

  We can write

$$\iint_S z\cos\gamma\, dS = \iint_{S_1} z\cos\gamma\, dS + \iint_{S_2} z\cos\gamma\, dS$$

Here, of the two integrals on the right hand side, the first is over the upper side of the upperhemisphere in the upward direction and the second is over the lower side of the lower hemisphere in the downward direction. The second integral will be negative and will therefore have to be taken with a negative sign.

$$\therefore \qquad \iint_{S_1} z \cos \gamma \, dS = \iint_{S_1} z \, dx \, dy = \iint_D \sqrt{1 - x^2 - y^2} \, dx \, dy$$

$$\iint_{S_2} z \cos \gamma \, dS = \iint_{S_1} z \, dx \, dy = -\iint_D -\sqrt{1 - x^2 - y^2} \, dx \, dy$$

where D is the circle $x^2 + y^2 = 1$.

Thus

$$\iint_S z \cos \gamma \, dS = 2 \iint_D \sqrt{1 - x^2 - y^2} \, dx \, dy = 2 \int_0^{2\pi} d\theta \int_0^1 \sqrt{1 - r^2} \, r \, dr = \frac{4\pi}{3}.$$

Note: The surface integral is the flux of a function through the surface S. In the above example, the flux through the whole surface is equal to the sum of the fluxes through the two hemisphere, ignoring the direction.

Example 12. Show that

$$I = \iint_S (yz \, dy \, dz + zx \, dz \, dx + xy \, dx \, dy) = \frac{3}{8}$$

where S is the other surface of the sphere $x^2 + y^2 + z^2 = 1$ in the first octant.

■ Let $\left.\begin{array}{l} x = \cos \theta \cos \phi \\ y = \cos \theta \sin \phi \\ z = \sin \theta \end{array}\right\}, \begin{array}{l} 0 \le \theta \le \pi/2 \\ 0 \le \phi \le \pi/2 \end{array}$

so that

$$\frac{\partial(y, z)}{\partial(\theta, \phi)} = -\cos^2 \theta \cos \phi, \frac{\partial(z, x)}{\partial(\theta, \phi)} = -\cos^2 \theta \sin \phi, \frac{\partial(x, y)}{\partial(\theta, \phi)} = -\sin \theta \cos \theta$$

the negative signs show that the correspondence is inverse and so the double integrals are to be taken with negative signs.

Thus we get

$$I = 3 \int_0^{\pi/2} \cos^3 \theta \sin \theta \, d\theta \int_0^{\pi/2} \sin \phi \cos \phi \, d\phi = \frac{3}{8}$$

Ex. Do the above example by the method explained in solved example 11.

Example 13. Evaluate $\iint_S x \, dS$, where S is the entire surface of the solid bounded by the cylinder $x^2 + y^2 = 1$ and the planes $z = 0$, $z = x + 2$.

■ As shown in the figure, the surface consists of three parts: S_1, the circular base in the xy-plane, S_2, the elliptic plane section, *i.e.*, part of the plane $z = x + 2$ inside the cylinder $x^2 + y^2 = 1$, and S_3, the lateral surface of the cylinder.

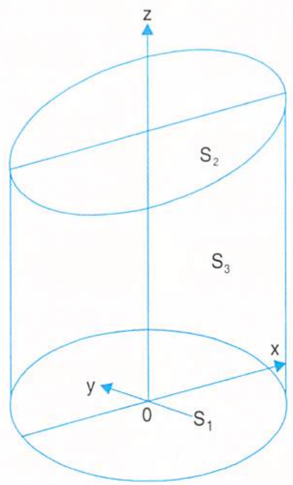


Fig. 6

On S_1, we have

$$z = 0, x^2 + y^2 = 1$$

$$\therefore \qquad \iint_{S_1} x\, dS = \iint_{S_1} x\, dx\, dy = \int_{-1}^{1} x\, dx \int_{-\sqrt{(1-x^2)}}^{\sqrt{(1-x^2)}} dy = 0$$

On S_2,

$$z = x + 2, x^2 + y^2 = 1$$

$$\therefore \qquad \iint_{S_2} x\, dS = \iint_{S_1} x\sqrt{1+1}\, dx\, dy = \sqrt{2} \iint_{S_1} x\, dx\, dy = 0$$

On S_2,

$$x^2 + y^2 = 1, 0 \le z \le x + 2$$

put

$$x = \cos\theta, y = \sin\theta, z = z$$
$$-\pi \le \theta \le \pi, 0 \le z \le 2 + \cos\theta$$

so that

$$\frac{\partial(y, z)}{\partial(\theta, z)} = \cos\theta, \frac{\partial(z, x)}{\partial(\theta, z)} = \sin\theta, \frac{\partial(x, y)}{\partial(\theta, z)} = 0$$

$$\therefore \qquad \iint_{S_2} x\, dS = \iint_D \cos\theta \sqrt{\cos^2\theta + \sin^2\theta}\, d\theta\, dz\,.$$

where D is the corresponding domain in the θz-plane.

$$= \int_{-\pi}^{\pi} \cos\theta \, d\theta \int_0^{2+\cos\theta} dz = \int_{-\pi}^{\pi} \cos\theta (2 + \cos\theta) \, d\theta = \pi$$

$$\iint_S x \, dS = \iint_{S_1} x \, dS + \iint_{S_2} x \, dS + \iint_{S_3} x \, dS = \pi \, .$$

Example 14. Evaluate $\iint_S \mathbf{F} \cdot \mathbf{n} \, dS$, where S is the entire surface of the solid formed by $x^2 + y^2 = 1, z = 0, z = x + 2$ and $\mathbf{n}$ is the outward drawn unit normal and the vector function $\mathbf{F} = 2x\mathbf{i} - 3y\mathbf{j} + z\mathbf{k}$.

■ The surface S consists of three sub-surfaces, the circular base S_1, the elliptic plane section S_2 and the lateral circular section S_3. (figure of example 13).

On S_1,

$$x^2 + y^2 = 1, z = 0$$

$$\mathbf{n} = -\mathbf{k}, \mathbf{F} \cdot \mathbf{n} = -z$$

$$\therefore \quad \iint_{S_1} \mathbf{F} \cdot \mathbf{n} \, dS = \iint_{S_1} -z \, dS = \iint_{S_1} -z \, dx \, dy = 0$$

On S_2,

$$z = x + 2, \; x^2 + y^2 = 1$$

$$-\frac{1}{\sqrt{2}} \mathbf{i} + 0\mathbf{j} + \frac{1}{\sqrt{2}} \mathbf{k}$$

$$\mathbf{F} \cdot \mathbf{n} = \frac{1}{\sqrt{2}} (-2x + z) = \frac{1}{\sqrt{2}} (-x + 2)$$

$$\therefore \quad \iint_{S_2} \mathbf{F} \cdot \mathbf{n} \, dS = \frac{1}{\sqrt{2}} \iint_{S_2} (-2x + z) \, dS$$

$$= \frac{1}{\sqrt{2}} \iint_{S_1} (-2x + x + 2) \sqrt{1 + 1} \, dx \, dy$$

$$= \iint_{S_1} (2 - x) \, dx \, dy$$

where S_1 is the circular base, $x^2 + y^2 = 1, z = 0$.

$$\iint_{S_2} F.n \, dS = \int_{-\pi}^{\pi} d\theta \int_0^1 (2 - r\cos\theta) \, r \, dr = \int_{-\pi}^{\pi} (1 - \frac{1}{3}\cos\theta) d\theta = 2\pi$$

On S_3,

$$x^2 + y^2 = 1, 0 \le z \le x + 2$$

Let

$$x = \cos\theta, \; y = \sin\theta, \; z = z$$

$$-\pi \le \theta \le \pi,\ 0 \le z \le 2 + \cos\theta$$

$$\mathbf{F}\cdot\mathbf{n} = (\mathbf{i}2\cos\theta - \mathbf{j}3\sin\theta + \mathbf{k}z)\cdot(\mathbf{i}\cos\theta + \mathbf{j}\sin\theta)$$

$$= 2\cos^2\theta - 3\sin^2\theta$$

$$\therefore \quad \iint_{S_3} \mathbf{F}\cdot\mathbf{n}\, dS$$

$$= \int_{-\pi}^{\pi}\int_0^{2+\cos\theta} (2\cos^2\theta - 3\sin^2\theta)\sqrt{\cos^2\theta + \sin^2\theta}\, d\theta\, dz$$

$$= \int_{-\pi}^{\pi} (2 - 5\sin^2\theta)(2 + \cos\theta) = -2\pi$$

$$\therefore \quad \iint_S \mathbf{F}\cdot\mathbf{n}\, dS = \iint_{S_1} \mathbf{F}\cdot\mathbf{n}\, dS + \iint_{S_2} \mathbf{F}\cdot\mathbf{n}\, dS + \iint_{S_3} \mathbf{F}\cdot\mathbf{n}\, dS$$

$$= 0 + 2\pi - 2\pi = 0\,.$$

EXERCISE

1. Evaluate the integrals

   (i) $\iint_S (z + 2x + \frac{4}{5}y)\, dS$, over the plane $\frac{x}{2} + \frac{y}{3} + \frac{z}{4} = 1$, lying in the first octant.

   (ii) $\iint_S xyz\, dS$, over the portion of $x + y + z = 1$, lying in the first octant.

   (iii) $\iint_S x\, dS$ where S is the portion of the sphere $x^2 + y^2 + z^2 = a^2$ lying in the first octant.

2. Compute the integrals:

   (i) $\iint_S \sqrt{a^2 - x^2 - y^2}\, dS$, where S is the hemisphere $z = \sqrt{a^2 - x^2 - y^2}$.

   (ii) $\iint_S x^2y^2\, dS$, where S is the hemisphere $z = \sqrt{R^2 - x^2 - y^2}$.

   (iii) $\iint_S \frac{dS}{r^2}$, where S is the cylinder $x^2 + y^2 = a^2$, bounded by the planes $z = 0$, $z = h$, and r is the distance between a point on the surface and the origin.

3. Evaluate the surface integrals:

   (i) $\iint_S x^2y^2 z\, dS$, where S is the positive side of the lower half of the sphere $x^2 + y^2 + z^2 = a^2$.

   (ii) $\iint_S z^2 dx\, dy$, where S is the outer side of the ellipsoid $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$.

   (iii) $\iint_S (x^3\, dy\, dz + y^3\, dz\, dx + z^3\, dx\, dy)$, where S is the outer surface of the sphere $x^2 + y^2 + z^2 = 1$.

(iv) $\iint_S (xz\,dx\,dy + xy\,dy\,dz + yz\,dz\,dx)$, where S is the outer side of the pyramid formed by the planes $x = 0$, $y = 0$, $z = 0$ and $x + y + z = 1$.

4. Evaluate the surface integral

$$\iint_S (x \cos \alpha + y \cos \beta + z \cos \gamma)\, dS.$$

$\cos \alpha$, $\cos \beta$, $\cos \gamma$ being the direction cosines of the outward drawn normal of the surface S, where

(i) S is the positive side of the cube formed by the planes $x = 0$, $y = 0$, $z = 0$, and $x = 1$, $y = 1$, $z = 1$.

(ii) S is the outer surface of the ellipsoid $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} + \dfrac{z^2}{c^2} = 1$ lying above the xy-plane.

5. Evaluate $\iint_E (x + y + z)(lx + my + nz)\, dS$, where E is the surface of the region $x^2 + y^2 \le a^2$, $0 \le z \le h$.

6. Find the value of the surface integral

$$\iint_S (yz\,dx\,dy + xz\,dy\,dz + xy\,dx\,dz),$$

where S is the outer side of the surface situated in the first octant and formed by the cylinder $x^2 + y^2 = a^2$ and the planes $x = 0$, $y = 0$, $z = 0$, $z = h$.

7. Evaluate $\iint_S z^2\, dS$, where S is the part of the outer surface of the cylinder $x^2 + y^2 = 4$ between the planes $z = 0$, $z = x + 3$.

8. Compute the integral

$$\iint_S (y^2 z\,dx\,dy + xz\,dy\,dz + x^2 y\,dx\,dz),$$

where S is the outer side of the surfaces situated in the first octant and formed by the paraboloid of revolution $z = x^2 + y^2$, cylinder $x^2 + y^2 = 1$ and the coordinate planes.

[See solved Example No. 24 for another method.]

9. Compute $\iint_S \mathbf{F} \cdot \mathbf{n}\, dS$, when $\mathbf{n}$ is the outward drawn unit normal of the surface, and the vector function $\mathbf{F}$ and the oriented surface S are given in each case.

(i) $\mathbf{F} = x^2\mathbf{i} + y^2\mathbf{j} + z^2\mathbf{k}$, S is part cone $z^2 = x^2 + y^2$, $1 \le z \le 2$, $\mathbf{n} \cdot \mathbf{k} > 0$, $\mathbf{n} \cdot \mathbf{i} > 0$.

(ii) $\mathbf{F} = y^2\mathbf{i} + z\mathbf{j} - x\mathbf{k}$, S is part of cylinder $y^2 = 1 - x$ between the planes $z = 0$, $z = x$; $x \ge 0$ with $\mathbf{n} \cdot \mathbf{i} > 0$.

ANSWERS

1. (i) $4\sqrt{61}$ (ii) $\sqrt{3}/120$ (iii) $\pi a^3/4$
2. (i) πa^3 (ii) $2\pi R^6/15$ (iii) $2\pi \tan^{-1}(h/a)$
3. (i) $2\pi a^7/105$ (ii) 0 (iii) $56\pi a^2/9$ (iv) 1/8
4. (i) 3 (ii) $2\pi abc$
5. $\frac{1}{3}\pi ah[3(l + m)a^2 + 3nah + 2nh^2]$
6. $a^2h(2a/3 + \pi h/8)$
7. π
8. $\pi/8$
9. (i) $15\pi/4$ (ii) 4/15

5. STOKES' THEOREM (*First generalization of Green's theorem*)

We recall that Green's Theorem expresses a relation between a double integral over a plane region and a line integral taken round its plane boundary. There are two ways to generalise this in R^3. One of these extensions, known as *Stokes' Theorem*, relates a surface integral taken over a surface to a line integral taken around the boundary curve of the surface. This generalisation is due to an English mathematician, *George Gabriel Stokes* (1819–1903).

A second generalisation arises when the double integral is replaced by a triple integral, and the line integral by a surface integral. This generalisation is named as *Gauss's Theorem* and will be taken up later.

Stokes' Theorem. *If S is a smooth oriented surface bounded by a curve C oriented in the same sense, and f, g, h are three functions which along with their first order partial derivatives are continuous in a three dimensional domain containing S, then*

$$\int_C (f\,dx + g\,dy + h\,dz) = \iint_S \left[\left(\frac{\partial h}{\partial y} - \frac{\partial g}{\partial z}\right) dy\,dz\right.$$

$$\left. + \left(\frac{\partial f}{\partial z} - \frac{\partial h}{\partial x}\right) dz\,dx + \left(\frac{\partial g}{\partial x} - \frac{\partial f}{\partial y}\right) dx\,dy\right]$$

Let the oriented surface be represented as

$$x = x(u, v),\ \ y = y(u, v),\ \ z = z(u, v),\ \ (u, v) \in D$$

where D is an oriented surface in uv-plane. Also, let its boundary be an oriented curve Γ represented by

$$u = u(t),\ \ v = v(t),\ \ a \le t \le b$$

The proof of the theorem involves the following steps:

1. The line integral along C is expressed as an ordinary integral,
2. The ordinary integral is expressed as a line integral along Γ,
3. The line integral along Γ is then expressed, by Green's Theorem, as double integral over D, and finally,
4. The double integral along D is expressed as a surface integral over S.

Now

$$\int_C (f\,dx + g\,dy + h\,dz)$$

$$= \int_a^b \left[f\left(\frac{\partial x}{\partial u}\frac{du}{dt} + \frac{\partial x}{\partial v}\frac{dv}{dt}\right) dt + g\left(\frac{\partial y}{\partial u}\frac{du}{dt} + \frac{\partial y}{\partial v}\frac{dv}{dt}\right) dt + h\left(\frac{\partial z}{\partial u}\frac{du}{dt} + \frac{\partial z}{\partial v}\frac{dv}{dt}\right) dt \right]$$

$$= \int_\Gamma \left(f\frac{\partial x}{\partial u} + g\frac{\partial y}{\partial u} + h\frac{\partial z}{\partial u}\right) du + \left(f\frac{\partial x}{\partial v} + g\frac{\partial y}{\partial v} + h\frac{\partial z}{\partial v}\right) dv$$

$$= \iint_D \left[\frac{\partial}{\partial u}\left(f\frac{\partial x}{\partial v} + g\frac{\partial y}{\partial v} + h\frac{\partial z}{\partial v}\right) - \frac{\partial}{\partial v}\left(f\frac{\partial x}{\partial u} + g\frac{\partial y}{\partial u} + h\frac{\partial z}{\partial u}\right)\right] du\,dv \qquad ...(1)$$

CHAPTER 18

But

$$\frac{\partial}{\partial u}\left(f\,\frac{\partial x}{\partial v}\right)=\left(\frac{\partial f}{\partial x}\frac{\partial x}{\partial u}+\frac{\partial f}{\partial y}\frac{\partial y}{\partial u}+\frac{\partial f}{\partial z}\frac{\partial z}{\partial u}\right)\frac{\partial x}{\partial v}+f\,\frac{\partial^2 x}{\partial u\partial v}$$

Writing down similar expressions for the other terms of the integrand and rearranging, the double integral on the right hand side of (1) becomes

$$\iint_D\left[\left(\frac{\partial h}{\partial y}-\frac{\partial g}{\partial z}\right)\frac{\partial(y,z)}{\partial(u,v)}+\left(\frac{\partial f}{\partial z}-\frac{\partial h}{\partial x}\right)\frac{\partial(z,x)}{\partial(u,v)}+\left(\frac{\partial g}{\partial x}-\frac{\partial f}{\partial y}\right)\frac{\partial(x,y)}{\partial(u,v)}\right]du\,dv$$

$$=\iint_S\left(\frac{\partial h}{\partial y}-\frac{\partial g}{\partial z}\right)dy\,dz+\left(\frac{\partial f}{\partial z}-\frac{\partial h}{\partial x}\right)dz\,dx+\left(\frac{\partial g}{\partial x}-\frac{\partial f}{\partial y}\right)dx\,dy \qquad \ldots(2)$$

Hence the proof.

Also by the definition of surface integral, relation (2) is equivalent to

$$\int_C (f\,dx+g\,dy+h\,dz)=\iint_S\left[\left(\frac{\partial h}{\partial y}-\frac{\partial g}{\partial z}\right)\cos\alpha+\left(\frac{\partial f}{\partial z}-\frac{\partial h}{\partial x}\right)\cos\beta+\left(\frac{\partial g}{\partial x}-\frac{\partial f}{\partial y}\right)\cos\gamma\right]dS \qquad \ldots(3)$$

where $\cos\alpha$, $\cos\beta$, $\cos\gamma$ are the direction cosines of the normal at any point to the surface.

Notes:

1. (*Vectorial formulation*). Let

$$\mathbf{r}=\mathbf{i}\,x+\mathbf{j}\,y+\mathbf{k}\,z$$

be the position vector of any point on the surface S, and

$$\mathbf{F}(x,y,z)=\mathbf{i}P(x,y,z)+\mathbf{j}Q(x,y,z)+\mathbf{k}R(x,y,z)$$

be a vector function defined on S.

Let $\mathbf{n}$ denote the unit normal at any point of the surface under consideration, so that

$$\mathbf{n}=\mathbf{i}\cos\alpha+\mathbf{j}\cos\beta+\mathbf{k}\cos\gamma$$

$$\therefore\quad \operatorname{curl}\mathbf{F}\cdot\mathbf{n}=\left(\frac{\partial R}{\partial y}-\frac{\partial Q}{\partial z}\right)\cos\alpha+\left(\frac{\partial P}{\partial z}-\frac{\partial R}{\partial x}\right)\cos\beta+\left(\frac{\partial Q}{\partial x}-\frac{\partial P}{\partial y}\right)\cos\gamma$$

and

$$\mathbf{F}\cdot d\mathbf{r}=P\,dx+Q\,dy+R\,dz$$

so that by (3) Stokes' theorem can be written as

$$\int_C \mathbf{F}\cdot d\mathbf{r}=\iint_S \operatorname{curl}\mathbf{F}.\mathbf{n}\,dS$$

2. If the surface S is a piece of a plane parallel to the xy-plane then $dz=0$ and we get Green's Theorem as special case of Stokes' Theorem.

5.1 Deductions from Stokes' Theorem

Stokes' theorem has various applications in mathematical analysis. Here we are going to establish only one such deduction: the conditions for a line integral to be independent of the path of integration. These conditions, in fact, generalise the results obtained from Green's theorem (§ 4.1 Ch. 17) concerning the question of path independence of an integral over a plane curve. With that view, we introduce the following concept.

Definition. A three-dimensional domain V is said to be *simply connected* if, for any closed contour belonging to V there exists a surface, with the contour as its boundary, entirely lying in V.

A sphere (ball), the whole space, the domain lying between two concentric spheres are examples of a *simply connected* space. An example of a domain which is not simply connected (*referred to as multiply connected*) is a ball with a cylindrical tunnel passing through it.

Now we proceed to establish the following result analogous to § 4.1, Chapter 17.

Theorem 2. *If three functions $f(x, y, z)$, $g(x, y, z)$, and $h(x, y, z)$, defined in a bounded closed simply connected domain V, are continuous along with their first order partial derivatives in the domain, then the following four assumptions are equivalent to each other.*

1. *The line integral $\int f\, dx + g\, dy + h\, dz$ taken along any closed contour lying inside V is equal to zero.*
2. *The line integral $\int_{AB} f\, dx + g\, dy + h\, dz$ is independent of the path of integration connecting two arbitrary points A and B.*
3. *The expression $f\, dx + g\, dy + h\, dz$ is the total differential of a single valued function defined in V.*
4. *The conditions*

$$\frac{\partial g}{\partial x} = \frac{\partial f}{\partial y}, \frac{\partial h}{\partial y} = \frac{\partial g}{\partial z}, \frac{\partial f}{\partial z} = \frac{\partial h}{\partial x}$$

are fulfilled at each point of the domain V.

The theorem is proved according to the scheme $1 \to 2 \to 3 \to 4 \to 1$ which we followed while proving § 4.1 Ch. 16. We leave the proof to the reader with the only hint that to deduce condition 1 from condition 4, one must take an arbitrary closed contour Γ lying within V and consider a surface entirely lying in V whose boundary is Γ, such a surface exists because of the condition that V is a simply connected domain. Then the application of Stoke's theorem to the line integral along Γ shows that condition 4 implies the relation $\int_\Gamma f\, dx + g\, dy + h\, dz = 0$.

Example 15. Use Stokes' theorem to find the line integral

$$\int_C x^2 y^3\, dx + dy + z\, dz$$

where C is the circle $x^2 + y^2 = a^2$, $z = 0$.

■ Now, by Stokes' theorem

$$\int_C x^2y^3\,dx + dy + z\,dz = \iint_S \left(\frac{\partial z}{\partial y} - \frac{\partial .1}{\partial z}\right) dy\,dz + \left(\frac{\partial x^2y^3}{\partial z} - \frac{\partial z}{\partial x}\right) dz\,dx$$

$$+ \left(\frac{\partial .1}{\partial x} - \frac{\partial x^2y^3}{\partial y}\right) dx\,dy = -\iint_S 3x^2y^2\,dx\,dy$$

where S is the circle $x^2 + y^2 = a^2$ in the xy-plane. Changing to polars.

$$= -3\int_{-\pi}^{\pi}\int_0^a r^4\cos^2\theta\sin^2\theta\, r\,dr\,d\theta$$

$$= -\frac{3a^6}{6}\int_{-\pi}^{\pi}\cos^2\theta\sin^2\theta\,d\theta = -\frac{\pi a^6}{8}$$

Example 16. Show that

$$\iint_S (y-z)\,dy\,dz + (z-x)\,dz\,dx + (x-y)\,dx\,dy = a^3\pi$$

where S is the portion of the surface $x^2 + y^2 - 2ax + az = 0,\ z \ge 0$.

■ By Stokes' theorem

$$\iint_S (y-z)\,dy\,dz + (z-x)\,dz\,dx + (x-y)\,dx\,dy$$

$$= \frac{1}{2}\int_C (y^2+z^2)\,dx + (z^2+x^2)\,dy + (x^2+y^2)\,dz$$

where C is the curve

$$(x-a)^2 + y^2 = a^2,\ \ z = 0$$

On putting $x = a + a\cos\theta = a(1+\cos\theta)$, $y = a\sin\theta$, the line integral becomes

$$= \frac{1}{2}\int_{-\pi}^{\pi}[a^2\sin^2\theta(-a\sin\theta) + a^2(1+\cos\theta)^2\,a\cos\theta]\,d\theta$$

$$= \frac{a^3}{2}\int_{-\pi}^{\pi}(-\sin^3\theta + \cos\theta + 2\cos^2\theta + \cos^3\theta)\,d\theta$$

$$= a^3\int_0^{\pi}(2\cos^2\theta + \cos^3\theta)\,d\theta$$

$$= a^3\int_0^{\pi/2}(2\cos^2\theta + \cos^3\theta)\,d\theta + a^3\int_0^{\pi/2}(2\sin^2\theta - \sin^3\theta)\,d\theta = \pi a^3.$$

Note: The method employed to convert a surface integral into a line integral is not general.

EXERCISE

1. Using Stokes' theorem, show that

$$\int_C y\,dx + z\,dy + x\,dz = -\iint_S(\cos\alpha + \cos\beta + \cos\gamma)\,dS.$$

2. Show, using Stokes' theorem, that

$$\int_\Gamma (y+z)\,dx + (z+x)\,dy + (x+y)\,dz = 0$$

where Γ is the circle $x^2 + y^2 + z^2 = a^2, x + y + z = 0$.

3. Using Stokes' theorem, prove that

$$\int_\Gamma y\,dx + z\,dy + x\,dz = -2\pi a^2\sqrt{2}$$

where Γ is the curve $x^2 + y^2 + z^2 - 2ax - 2ay = 0, x + y = 2a$.

4. Apply Stokes' theorem to transform the integral

$$\int_C (y^2 + z^2)\,dx + (x^2 + z^2)\,dy + (x^2 + y^2)\,dz$$

taken along a smooth curve C to a certain integral over a smooth oriented surface with C as its boundary.

5. Verify Stokes' theorem for the integral

$$\int_C x^2\,dx + yx\,dy$$

where C is a square in the $z = 0$ plane with sides along the lines, $x = 0, y = 0, x = a, y = a$.

6. Verify Stokes' theorem in each case

(i) $\mathbf{F} = z\mathbf{i} + x\mathbf{j} + y\mathbf{k}$

S is the part of the paraboloid $z = 1 - x^2 - y^2$ for which $z \geq 0$, $\mathbf{n}\cdot\mathbf{k} > 0$.

(ii) $\mathbf{F} = y^2\mathbf{i} + xy\mathbf{j} - 2xz\mathbf{k}$

S is the hemisphere $x^2 + y^2 + z^2 = a^2, z \geq 0$ with $n\cdot\mathbf{k} > 0$.

6. THE VOLUME OF A CYLINDRICAL SOLID BY DOUBLE INTEGRALS

We have shown earlier that the volume of a cylindrical solid can be found with the help of double integrals.

Let a cylindrical solid be bounded above by a surface $z = \Psi(x, y)$, below by a plane region D (on the xy-plane) and on the sides by lines parallel to z-axis. Its volume V is given by

$$V = \iint_D \Psi(x, y)\,dx\,dy, \text{ in cartesian coordinates}$$

or

$$= \iint_D \Psi(r\cos\theta, r\sin\theta)\,r\,dr\,d\theta, \text{ in polar coordinates}$$

or

$$= \iint_S z\cos\gamma\,dS, \text{ as a surface integral}$$

where S is the surface of the solid.

If the equation of the surface is given in the form

$$x = \theta(y, z), \text{ or } y = \phi(z, x)$$

then the corresponding formulas for calculating the volumes are of the form

$$V = \iint_{D_1} \theta(y, z)\, dy\, dz \quad \text{or} \quad \iint_{S_1} x \cos \alpha \, dS$$

$$V = \iint_{D_2} \phi(z, x)\, dz\, dx \quad \text{or} \quad \iint_{S_2} y \cos \beta \, dS$$

where D_1, D_2 are the domains in the yz-plane and zx-plane, in which the given surface is projected.

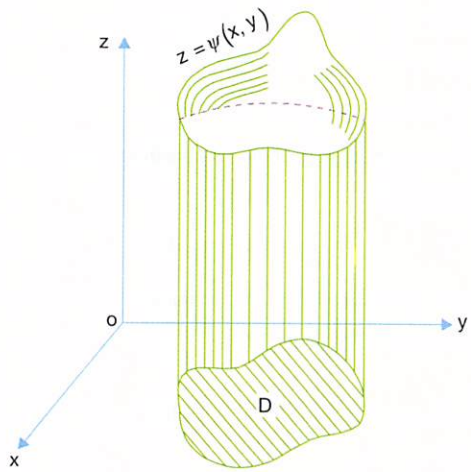


Fig. 7

Notes:

1. Clearly D is projection on the xy-plane, of the portion of the surface $z = \Psi(x, y)$, cut off by the lateral cylindrical surface.
2. The function ψ is assumed to be continuous and single-valued so that the surface is met by a line parallel to z-axis not more than one point.
3. If the function ψ changes sign in D, then we divide the domain D into two parts. (*i*) the sub-domain D_1 where $\Psi \geq 0$, and (*ii*) the sub-domain D_2 where $\Psi \leq 0$. The double integral over D_1 will be positive and equal to the volume of the solid lying above the xy-plane. The integral over D_2 will be negative and equal, in absolute value, to the volume of the solid lying below the xy-plane. Thus the integral over D will express the difference between the corresponding volumes. The sum of the absolute values of the two integrals, over D_1 and D_2, will give the volume of the solid.
4. *Volume by iterated integral interpreted geometrically.*

$$V = \iint_D \Psi\, dxdy = \int_a^b dx \int_{\phi_1(x)}^{\phi_2(x)} \Psi \, dy$$

$$= \int_a^b S(x)\, dx\,.$$

where

$$S(x) = \int_{\phi_1(x)}^{\phi_2(x)} \Psi \; dy = \text{area of a cross section parallel to } zx\text{-plane}$$

$\therefore \quad V =$ volume of the solid by parallel cross-sections

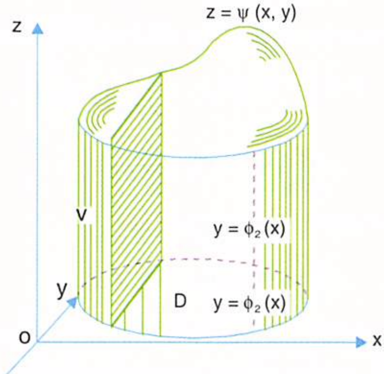


Fig. 8

6.1 Volume Enclosed by Two Surfaces

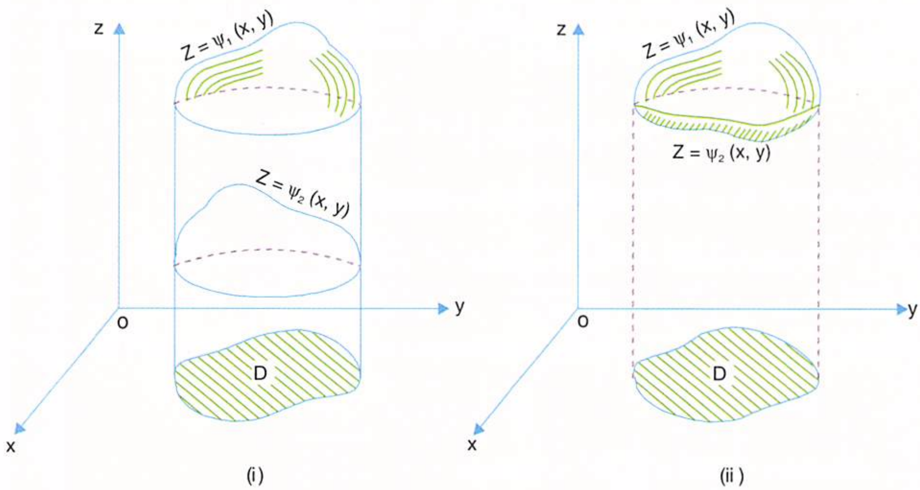


Fig. 9

If a solid, the volume of which is being found, is bounded above by the surface $z = \Psi_1(x, y) \geq 0$ and below by the surface $z = \Psi_2(x, y) \geq 0$, and domain D is the projection of both the surfaces (Fig. (*i*) and (*ii*)) on the xy-plane, then the volume, V, of this solid is equal to the difference between the volumes of the two 'cylindrical bodies'; the first of these cylindrical bodies has the domain D for its lower base and the surface $z = \Psi_1(x, y)$ for the upper; the second body also has D for its lower base and the surface $z = \Psi_2(x, y)$ for its upper base.

Therefore the required volume V is equal to the difference between the two double integrals.

$$\therefore \qquad V = \iint_D \Psi_1 \, dx \, dy - \iint_D \Psi_2 \, dx \, dy$$

$$= \iint_D (\Psi_1 - \Psi_2) \, dx \, dy$$

which may be expressed in terms of a surface integral as

$$= \iint_S (\Psi_1 - \Psi_2) \cos \gamma \, dS$$

It may be easily verified that the formula holds true not only for the case where Ψ_1 and Ψ_2 are non-negative, but also where Ψ_1 and Ψ_2 are any continuous, single-valued function that satisfy the relationship

$$\Psi_1(x, y) \geq \Psi_2(x, y) \text{ over } D$$

6.2 Volume Enclosed by a Closed Surface

Let a closed surface S be such that any straight line parallel to the z-axis cut it in not more than two points. Let the outer normal be the positive direction of the normal.

The surface may be divided into two parts: the upper and the lower. Let their equations be $z = \Psi_1(x, y)$, $z = \Psi_2(x, y)$. If D is the projection of S on xy-plane, then since the normal to the lower surface is downward, we have

$$\iint_S z \cos \gamma \, dS = \iint_D (\Psi_1 - \Psi_2) \, dx \, dy$$

which represents the volume of the solid under consideration.

Example 17. Find the volume within the cylinder $x^2 + y^2 = a^2$ between the planes $y + z = b^2$, and $z = 0$.

■ The cylindrical solid is bounded above by the surface $z = b^2 - y \equiv f(x, y)$, and below by the disc

$$D \equiv x^2 + y^2 = a^2.$$

$$\therefore \qquad \text{Required volume} = \iint_D (b^2 - y) \, dx \, dy$$

$$= \int_{-a}^{a} dx \int_{-\sqrt{(a^2 - x^2)}}^{\sqrt{(a^2 - x^2)}} (b^2 - y) \, dy = \pi a^2 b^2 .$$

Example 18. Find the volume of

(*i*) the solid bounded by the surface $z = 1 - 4x^2 - y^2$ and the plane $z = 0$.

(*ii*) the sphere $x^2 + y^2 + z^2 = a^2$, using polar coordinates.

■ (*i*) The solid in question is a segment of the elliptical paraboloid lying above the *xy*-plane. The paraboloid cuts the *xy*-plane along the ellipse $4x^2 + y^2 = 1$, which forms the base D of the solid. Thus the solid, without lateral cylindrical surface, is bounded above by $z = 1 - 4x^2 - y^2$ and below by the ellipse $4x^2 + y^2 = 1$. Moreover the solid being symmetrical, its volume V is four times the volume lying in the first octant.

$$\therefore \quad V = 4\int_0^{1/2} dx \int_0^{\sqrt{1-4x^2}} (1 - 4x^2 - y^2)\, dy = \frac{8}{3}\int_0^{1/2} (1 - 4x^2)^{3/2}\, dx$$

$$= \frac{4}{3}\int_0^{\pi/2} \cos^4 t\, dt = \frac{\pi}{4}.$$

(*ii*) As in Part (*i*) the solid under consideration is bounded above by $z^2 = a^2 - (x^2 + y^2)$, or by $z = \sqrt{a^2 - r^2}$, on changing to cylindrical polar coordinates. The sphere cuts the *xy*-plane in the circle $x^2 + y^2 = a^2$ or $r^2 = a^2$ and has no lateral cylindrical surface. Again because of symmetry, its volume

$$V = 2\int_0^{2\pi} d\theta \int_0^a \sqrt{a^2 - r^2}\, r\, dr = \frac{4}{3}\pi a^3.$$

Example 19. Compute the volume V, common to the ellipsoid of revolution $x^2/a^2 + y^2/a^2 + z^2/b^2 = 1$ and the cylinder $x^2 + y^2 - ay = 0$.

■ The required volume is double the volume that lies above the *xy*-plane.

The solid under consideration is bounded above by $z = \frac{b}{a}\sqrt{a^2 - x^2 - y^2}$ and below by the circular base $D \equiv x^2 + y^2 - ay = 0$ on the *xy*-plane.

In polar coordinates, the upper boundary is $z = \frac{b}{a}\sqrt{a^2 - r^2}$, and the lower base is $r = a\sin\theta$. Thus,

$$V = 2\int_0^{\pi/2} d\theta \int_0^{a\sin\theta} \frac{b}{a}\sqrt{a^2 - r^2}\, r\, dr$$

$$= \frac{4a^2b}{3}\int_0^{\pi/2} (1 - \cos^3\theta)\, d\theta = \frac{2}{9}a^2b(3\pi - 4).$$

Example 20. Compute the volume of the solid bounded by the cylindrical surfaces $z = 4 - y^2$ and $y = x^2/2$ and the plane $z = 0$.

The upper boundary of the solid is the surface with equation $z = 4 - y^2$. The domain of integration D is bounded by the parabola $y = x^2/2$ and the line of intersection of the cylinder $z = 4 - y^2$ with $z = 0$ plane, *i.e.*, the straight line $y = 2$.

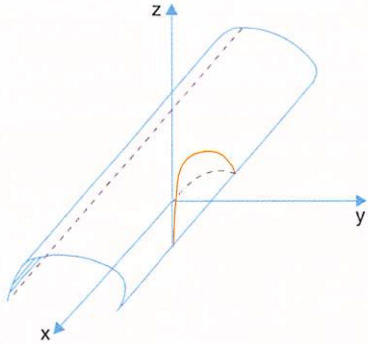


Fig. 10

The solid being symmetrical about the yz-plane, its volume

$$V = 2\int_0^2 dx \int_{x^2/2}^2 (4 - y^2)\, dy$$

$$= 2\int_0^2 \left(8 - \frac{8}{3} - 2x^2 + \frac{x^6}{24}\right) dx = \frac{256}{21}.$$

Example 21. Find the volume of the solid bounded above by the parabolic cylinder $z = 4 - y^2$ and bounded below by the elliptic paraboloid $z = x^2 + 3y^2$.

■ The two surfaces intersect in a space curve, whose projection on the xy-plane is the ellipse

$$x^2 + 4y^2 = 4$$

or

$$x^2/4 + y^2 = 1$$

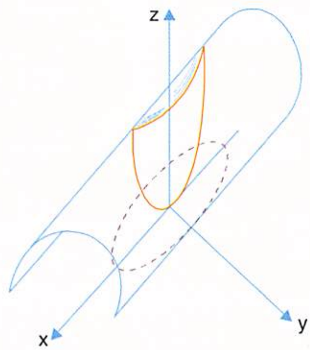


Fig. 11

which on putting

$$x = 2r\cos\theta,\ \ y = r\sin\theta$$

becomes $r^2 = 1$.

The difference,

$$f_1(x, y) - f_2(x, y) = 4 - 4y^2 - x^2 = 4(1 - r^2)$$

The Jacobian $J = 2r$.

Making use of the symmetry of the solid, the required volume

$$= 4\int_0^{\pi/2}\int_0^1 4(1 - r^2)2r\,dr\,d\theta = 4\pi\,.$$

EXERCISE

1. Show that the volume of the solid bounded above by $z = 2 - r$ and below by the plane region, $0 \le r \le 2\cos\theta, -\pi/2 \le \theta \le \pi/2$ is $2(9\pi - 16)/9$.
2. Prove that the volume common to the sphere $x^2 + y^2 + z^2 = a^2$ and the cylinder $x^2 + y^2 = ay$ is $2(3\pi - 4)a^3/9$.
3. Show that the volume common to the surface $y^2 + z^2 = 4ax$ and $x^2 + y^2 = 2ax$ is $\frac{2}{3}(3\pi + 8)a^3$.
4. Show that the volume common to the cylinders $x^2 + y^2 = a^2$ and $x^2 + z^2 = a^2$ is $16a^3/3$.
5. A sphere of radius a is pierced by a circular cylinder of radius b $(b < a)$, the axis of the cylinder passing through the centre of the sphere. Prove that the volume of the sphere that lies inside the cylinder is $\frac{4}{3}\pi\,[a^3 - (a^2 - b^2)^{3/2}]$ and the surface of the sphere inside the cylinder is $4\pi a\{a - (a^2 - b^2)^{1/2}\}$.
6. The part of the volume of a sphere of radius a that lies inside a right circular cone of semi-vertical angle α whose vertex is on the sphere and whose axis is a diameter of the sphere is $\frac{4}{3}\pi a^3\,(1 - \cos^4\alpha)$.
7. One loop of the curve $r^2\cos^2\theta = a^2\cos 2\theta$ makes a complete revolution about the initial line; show that the volume of the solid generated is $\dfrac{\pi}{6}(10 - 3\pi)a^3$.
8. If V is the volume of a solid bounded by a surface σ then show that
$$\iint_\sigma (x\cos\alpha + y\cos\beta + z\cos\gamma)\,d\sigma = 3V\,.$$
9. Find $\iint_S z\,dx\,dy$, where S is the external surface of the sphere $x^2 + y^2 + z^2 = a^2$.
10. The sphere $x^2 + y^2 + z^2 = a^2$ is pierced by the cylinder $(x^2 + y^2)^2 = a^2(x^2 - y^2)$, prove that the volume of the sphere inside the cylinder is $\dfrac{8}{3}\left(\dfrac{\pi}{4} + \dfrac{5}{3} - \dfrac{4\sqrt{2}}{3}\right)a^3$, and that the area of the spherical surface inside the cylinder is $8\left(\dfrac{\pi}{4} + 1 - \sqrt{2}\right)a^2$.

7. VOLUME INTEGRALS (*Triple Integrals*)

In the foregoing chapter we introduced the notion of the double integral. Here we are going to define the integral of a function of three independent variables, the so-called *volume integral*, also known as *triple integral*, in R^3.

Triple integrals are a straight and simple extension of the idea of double integrals and are in many respects almost completely analogous to them. We shall, therefore, only briefly indicate the various stages of the development of the theory and omit those proofs which do not essentially differ from the corresponding proofs of the theory of the double integrals.

7.1 Partition of a Rectangular Parallelepiped

A region in R^3, enclosed by the inequalities (including its boundary),

$$a \le x \le b; \quad c \le y \le d; \quad g \le z \le h$$

and denoted by $R = [a, b; c, d; g, h]$ is called a *rectangular parallelepiped* (a cuboid), or a *rectangle* in R^3. Its volume $V = (b-a)(d-c)(h-g)$.

If $P_1 = \{a = x_0, \ x_1, \ldots, x_l = b\}$, $P_2 = \{c = y_0, \ y_1, \ldots, y_m = d\}$, $P_3 = \{g = z_0, \ z_1, \ldots, z_n = h\}$ are respectively the partitions of $[a, b]$, $[c, d]$ and $[g, h]$, then planes drawn parallel to the coordinate planes through the points of P_1, P_2, P_3 give rise to a *partition* P of the parallelepiped R into lmn sub-parallelepipeds $[x_{i-1}, x_i; y_{j-1}, y_j; z_{k-1}, z_k]$ denoted as ΔR_{ijk}. The symbol ΔR_{ijk} will denote the sub-parallelepiped as also its volume $(x_i - x_{i-1})(y_j - y_{j-1})(z_k - z_{k-1})$.

Clearly $P = P_1 \times P_2 \times P_3$, the Cartesian product of P_1, P_2, P_3.

If $\mu(P_1) = \Delta x_r$, $\mu(P_2) = \Delta y_s$, $\mu(P_3) = \Delta z_t$ be the norms of P_1, P_2, P_3 respectively, then $\Delta R_{rst} = [x_{r-1}, x_r; y_{s-1}, y_s; z_{t-1}, z_t]$ is called the *norm* of P, denoted as $\mu(P)$.

Clearly, the volume of each sub-parallelepiped tends to zero as $\mu(P)$ tends to zero.

7.2 Triple Integration over a Parallelepiped

Let f be a bounded function of x, y, z on a parallelepiped $R = [a, b; c, d; g, h]$, and P, a partition of R. As in double integrals, we form the *Darboux sums*

$$U(P, f) = \sum_{i=1}^{l} \sum_{j=1}^{m} \sum_{k=1}^{n} M_{ijk} \, \Delta R_{ijk}, \text{ the } \textit{upper sum}$$

$$L(P, f) = \sum_{i=1}^{l} \sum_{j=1}^{m} \sum_{k=1}^{n} m_{ijk} \, \Delta R_{ijk}, \text{ the } \textit{lower sum}$$

where m_{ijk}, M_{ijk} are the lower and the upper bounds of f in ΔR_{ijk}.

It may be easily shown that

$$m(b-a)(d-c)(h-g) \le L(P, f) \le U(P, f) \le M(b-a)(d-c)(h-g)$$

The infimum of the set of upper sums and the supremum of the set of lower sums, for all partitions of R, are respectively known as the *upper* and the *lower integrals* of f over R, and are denoted as

$$\overline{\iiint}_R f\,dx\,dy\,dz \quad \text{and} \quad \underline{\iiint}_R f\,dx\,dy\,dz$$

In case the two integrals are equal, f is said to be integrable—the common value denoted by

$$\iiint_R f\,dx\,dy\,dz \quad \text{or} \quad \int_a^b \int_c^d \int_g^h f\,dx\,dy\,dz$$

is called the *triple integral* of f over R.

7.3 The following *necessary and sufficient* condition for the existence of the triple integral is proved by applying arguments similar to those for the double and single integrals.

Theorem 3. *A bounded function f defined on a parallelepiped R is integrable on R if and if for every $\varepsilon > 0$ there is a partition P of R such that*

$$U(P, f) - L(P, f) < \varepsilon$$

The criterion implies all theorems (including Darboux theorem) similar to those of § 2.6 and 2.7 for double integrals. Three of the theorems may be stated as follows:

1. **Darboux Theorem.** *If f is a bounded function on R then to every $\varepsilon > 0$ there corresponds $\delta > 0$ such that*

   $$U(P, f) < I^u + \varepsilon, \;\; L(P, f) > I_l - \varepsilon$$

   for every partition P of R with norm $\mu(P) < \delta$.
2. *Every continuous function is integrable.*
3. *A bounded function is integrable if its discontinuities are finite in number or if, when infinite in number, they all lie on a finite number of surfaces, which can therefore be enclosed in a finite number of parallelepipeds whose total volume can be made arbitrarily small.*

7.4 Reduction to Iterated Integrals
(Calculation of a triple integral over a parallelepiped)

At iterated integrals is an integral of the form

$$\iint_{R_1} dx\,dy \int_g^h f(x, y, z)\,dz \quad \text{or} \quad \iint_{R_1} \left[\int_g^h f(x, y, z)\,dz\right] dx\,dy$$

where $R_1 = [a, b; c, d]$ is the projection of $R = [a, b; c, d; g, h]$ on the xy-plane.

The proof of the theorem relating to the reduction of a triple integral to iterated integrals and is similar to that for double integrals will therefore be only briefly indicated.

Theorem 4. *If*

(*i*) *the triple integral $\iiint_R f(x, y, z)\,dx\,dy\,dz$ exists over $R = [a, b; c, d; g, h]$,*

(*ii*) *the integral $\Psi(x, y) = \int_g^h f(x, y, z)\,dz$ exists for every fixed point (x, y) of $R_1 = [a, b; c, d]$,*

then the integral $\iint_{R_1} \left[\int_g^h f(x, y, z)\,dz\right] dx\,dy$ also exists, and

$$\iiint_R f(x, y, z)\, dx\, dy\, dz = \iint_{R_1} \left[\int_g^h f(x, y, z)\, dz \right] dx\, dy \qquad ...(1)$$

Let I^u and I_l denote the upper and the lower triple integrals of f over R.

Let $\varepsilon > 0$ be an arbitrary number.

There exists a partition P of R such that (with usual notation)

$$\sum_i \sum_j \sum_k M_{ijk}\, \Delta x_i\, \Delta y_j\, \Delta z_k < I^u + \varepsilon$$

For each fixed (x, y) in $R_1 = [a, b; c, d]$ we have

$$\overline{\iint}_{R_1} \left[\overline{\int}_g^h f(x, y, z)\, dz \right] dx\, dy \le \overline{\iint}_{R_1} \left[\sum_k M_{ijk}\, \Delta z_k \right]$$

$$\le \sum_i \sum_j \left(\sum_k M_{ijk}\, \Delta z_k \right) \Delta x_i\, \Delta y_i < I^u + \varepsilon$$

But by hypothesis, $\overline{\int}_c^d f(x, y, z)\, dz = \int_c^d f(x, y, z)\, dz$, and since ε is an arbitrary positive number,

$$\therefore \qquad \overline{\iint}_{R_1} \left[\int_g^h f(x, y, z)\, dz \right] dx\, dy \le I^u + \varepsilon \qquad ...(2)$$

$$\le I^u = \overline{\iiint}_R f(x, y, z)\, dx\, dy\, dz$$

It can be similarly shown that

$$\underline{\iint}_{R_1} \left[\int_g^h f(x, y, z)\, dz \right] dx\, dy \ge I_l = \underline{\iiint}_R f(x, y, z)\, dx\, dy\, dz \qquad ...(3)$$

Again, since $I^u = I_l = I$, as the triple integral exists, we get from equations (2) and (3)

$$I \le \underline{\iint}_{R_1} \left[\int_g^h f(x, y, z)\, dz \right] dx\, dy \le \overline{\iint}_{R_1} \left[\int_g^h f(x, y, z)\, dz \right] dx\, dy \le I$$

$$\Rightarrow \qquad \underline{\iint}_{R_1} \left[\int_g^h f(x, y, z)\, dz \right] dx\, dy = \overline{\iint}_{R_1} \left[\int_g^h f(x, y, z)\, dz \right] dx\, dy = I$$

$$\Rightarrow \qquad \iint_{R_1} \left[\int_g^h f(x, y, z)\, dz \right] dx\, dy \text{ exists and equals } \iiint_R f(x, y, z)\, dx\, dy\, dz$$

Hence the theorem.

Corollary 1. If, further we assume that the double integral can be reduced to iterated integrals (*i.e.*, $\phi(x) = \int_c^d \Psi(x, y)\, dy$ exists for each fixed x in $[a, b]$), then

$$\iint_{R_1} \Psi(x, y)\, dx\, dy = \int_a^b dx \int_c^d \Psi(x, y)\, dy$$

$$= \int_a^b dx \int_c^d \left[\int_g^h f(x, y, z)\, dz\right] dy$$

and so we deduce that

$$\iiint_R f(x, y, z)\, dx\, dy\, dz = \int_a^b dx \int_c^d dy \int_g^h f(x, y, z)\, dz \qquad ...(4)$$

The formula reduces the evaluation of a triple integral over a parallelepiped R to successive separate integrations with respect to each variable. The integration is performed first with respect to z, then with respect to y and finally with respect to x.

Similarly, if the integrals

$$\Psi_1(y, z) = \int_a^b f(x, y, z)\, dx \quad \text{and} \quad \int_c^d \Psi_1(y, z)\, dy$$

exist, then we derive the analogous formula

$$\iiint_R f(x, y, z)\, dx\, dy\, dz = \int_g^h dz \int_c^d dy \int_a^b f(x, y, z)\, dx \qquad ...(5)$$

Similarly, on condition that the corresponding single and double integrals exist, we can establish analogous formulas reducing the triple integral to iterated integrals with respect to x, y and z in various orders.

Corollary 2. In particular, if the function f is continuous, the triple and all the possible double and single integrals are sure to exist and therefore the triple integral can be evaluated by expressing it as an iterated integral of x, y, z in any order.

However, in the general case of an arbitrary integrable function f, the orders are not always interchangeable.

7.5 Triple Integral over Regions (*Bounded domains*)

Definition 1. Let a bounded function f (of three variables) be defined on a region E of volume V.

Let a partition, consisting of a finite number of surfaces, divide the region E into n sub-regions of elementary volumes $\Delta V_1, \Delta V_2, \ldots, \Delta V_n$.

Clearly the sum of these volumes is $\sum \Delta V_i = V$.

With, usual notation form the sums

$$U(P, f) = \sum_i M_i\, \Delta V_i, \quad \text{and} \quad L(P, f) = \sum_i m_i\, \Delta V_i$$

respectively called the upper and the lower (Darboux) sums. It can be easily shown that

$$mV \le L(P, f) \le U(P, f) \le MV$$

so that the two sets of sums, the upper and the lower sum (corresponding to all the partitions of E) are bounded.

The infimum of the set of upper sums, and the supremum of the set of lower sums are respectively called the *upper* and the *lower integrals* denoted as

$$I^u = \overline{\iiint}_E f\ dV \quad \text{or} \quad \overline{\iiint}_E f\ dx\ dy\ dz$$

and

$$I_l = \underline{\iiint}_E f\ dV \quad \text{or} \quad \underline{\iiint}_E f\ dx\ dy\ dz$$

When these two integrals are equal, f is said to be integrable and the common value, called the *integral* of f over E, is denoted as

$$I = \iiint_E f\ dV \quad \text{or} \quad \iiint_E f\ dx\ dy\ dz \qquad \text{...(1)}$$

Remark: (*Volume of a solid*). Taking $f = 1$ in (1) we deduce that the volume enclosed by a closed region E (of any shape) is given by

$$V = \iiint_E dx\ dy\ dz \qquad \text{...(2)}$$

Definition 2. (*Integral as a limit of the sums*). If (ξ_i, η_i, ζ_i) is any point of ΔV_i, then the limit

$$\lim_{\mu(P)\to 0} S(P, f) = \lim_{\mu(P)\to 0} \sum_i f(\xi_i, \eta_i, \zeta_i)\ \Delta V_i$$

if it exists, for all partitions P of E and for all positions of the point (ξ_i, η_i, ζ_i) in ΔV_i, is called the triple integral of f over E. Thus

$$\lim_{\mu(P)\to 0} S(P, f) = \iiint_E f\ dV = \iiint_E f\ dx\ dy\ dz \qquad \text{...(3)}$$

7.6 Some Theorems

The basic properties of triple integrals are completely analogous to those of the double integrals and as such all theorems of double integrals mentioned in § 2.3, 2.6, 2.7 of Chapter 17 and those of triple integrals over parallelepipeds (§ 7.3) hold good for triple integrals over any region E.

We may enumerate some of them here:

1. Darboux Theorem.
2. The necessary and sufficient conditions for integrability.

   The statements and proofs are exactly same as for triple integrals over parallelepiped except that R is replaced by E.

   If functions f_1, f_2 and f_3 of x, y, z are integrable over a region E in R^3, then
3. (Linearity)

$$\iiint_E (k_i f_1 + k_2 f_2)\ dx\ dy\ dz = k_1 \iiint_E f_1\ dx\ dy\ dz + k_2 \iiint_E f_2\ dx\ dy\ dz$$

   where k_1, k_2 are constants.

4. (Additivity). If E is union of two regions E_1, E_2 with no interior points in common and f is integrable over E_1 and E_2 then f is integrable over E and

$$\iiint_E f\ dx\ dy\ dz = \iiint_{E_1} f\ dx\ dy\ dz + \iiint_{E_2} f\ dx\ dy\ dz\,.$$

5. (Monotonicity). If $f_1(x, y, z) \geq f_2(x, y, z)$, then

$$\iiint_E f_1\ dx\ dy\ dz \geq \iiint_E f_2\ dx\ dy\ dz\,.$$

6. When f is integrable over E, so is $|f|$ and

$$\left|\iiint_E f\ dx\ dy\ dz\right| \leq \iiint_E |f|\ dx\ dy\ dz\,.$$

7. (Mean Value Theorems). If $m \leq f(x, y, z) \leq M$, and V is the volume of E, then

$$mV \leq \iiint_E f\ dx\ dy\ dz \leq MV\,.$$

   If f is continuous, then there exists a point (ξ, η, ζ) of E such that

$$\iiint_E f\ dx\ dy\ dz = V \cdot f(\xi, \eta, \zeta)\,.$$

8. Every continuous function is integrable.
9. A bounded function is integrable if its discontinuities are finite in number or if, when infinite in number, they all lie on a finite number of surfaces, which can therefore be enclosed in a finite number of volumes whose total volume can be made arbitrarily small.

CHAPTER 18

7.7 Volume of Solids by Triple Integrals

Volume V of a solid of any shape in R^3, (§ 7.5) is given by

$$V = \iiint_E dx\ dy\ dz \qquad \ldots(1)$$

It may be noted that whereas the formulas of § 6 are helpful to find the volume of *cylindrical* solids only, the present formula enables us to find the volume of a solid of *any shape* in R^3.

7.8 Regular (or Quadratic) Domain

Definition. A three-dimensional domain is called regular (or quadratic) with respect to z-axis if every straight line parallel to the z-axis and passing through a point (not lying on the boundary) of the domain, cuts its surface in not more than two points.

A domain E bounded above and below by the surfaces

$$z = \phi(x, y),\ \ z = \Psi(x, y),\ \ (\phi(x, y) \geq \Psi(x, y))$$

and on the sides by a lateral cylindrical surface, is *regular with respect to z-axis.*

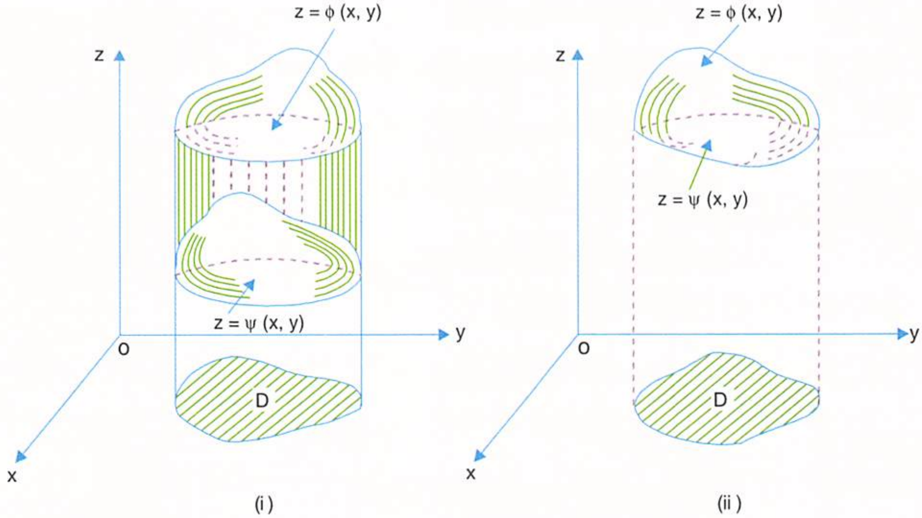


Fig. 12

Such domains are projectable on the xy-plane into a regular (two dimensional) domain D. Further any part of such a domain cut off by a plane parallel to any of the coordinate planes is also regular with respect to z-axis.

Domains regular with respect to the axis of x or y are similarly defined.

A domain which is regular with respect to all the axes is called a *regular domain.*

A domain which can be divided into a finite number of domains each of which is regular with respect to an axis is called *piecewise regular* with respect to that axis.

It may be noted that the definition also includes the cases where there can be no lateral surface. For instance, a three dimensional sphere (ball) is considered to be a domain regular with respect to z-axis (all the axes) which is bounded above and below by the upper and the lower hemispheres and whose lateral surface degenerates into the equator.

7.9 Reduction to Iterated Integrals (*Calculation of a triple integral over any region in R^3*)

Theorem 5. *If a triple integral* $\iiint_E f\ dx\ dy\ dz$ *exists for a function f defined on a closed* (*regular*) *domain E bounded above and below by the surfaces*

$$z = \phi(x, y),\ \ z = \Psi(x, y),\ \ (\Psi(x, y) \geq \phi(x, y))$$

and on the sides by a cylindrical surface, and if the integral $\int_{\phi(x,y)}^{\Psi(x,y)} f(x, y, z)\ dz$ *exists for each fixed point* (x, y) *belonging to D* (*the projection of E on xy-plane*), *then the iterated integral* $\iint_D \left[\int_{\phi(x,y)}^{\Psi(x,y)} f(x, y, z)\ dz\right] dx\ dy$ *also exists and*

$$\iiint_E f(x, y, z)\, dx\, dy\, dz = \iint_D \left[\int_{\phi(x,y)}^{\Psi(x,y)} f(x, y, z)\, dz \right] dx\, dy$$

Let a parallelepiped $R = [a, b; c, d; g, h]$ encloses the bounded domain E. Then obviously the projection R_1 of R on xy-plane encloses D.

Let us define a function F over R such that

$$F(x, y, z) = \begin{cases} f(x, y, z), & \text{at points of } E \\ 0 & \text{outside } E \end{cases}$$

Then

$$\iiint_E f(x, y, z)\, dx\, dy\, dz = \iiint_R F(x, y, z)\, dx\, dy\, dz$$

$$\iint_{R_1} \left[\int_g^h F(x, y, z)\, dz \right] dx\, dy \qquad \text{...(1)}$$

The function $F(x, y, z)$ vanishes outside E, therefore we have

$$\int_g^h F(x, y, z)\, dz = \int_{\phi(x,y)}^{\Psi(x,y)} f(x, y, z)\, dz \qquad \text{...(2)}$$

The expression (2) is a function of x and y which is equal to zero outside the domain D. Therefore, the double integral of the expression taken over R_1 coincides with its double integral over D. Hence (1) and (2) give

$$\iiint_E f(x, y, z)\, dx\, dy\, dz = \iint_D \left[\int_{\phi(x,y)}^{\Psi(x,y)} f(x, y, z)\, dz \right] dx\, dy \qquad \text{...(3)}$$

Hence the theorem.

Remarks:

1. If further, the conditions for the reduction of the double integral to the iterated integrals are satisfied, *i.e.*, if

$$\iint_D I(x, y)\, dx\, dy = \int_a^b dx \int_{\phi_1(x)}^{\Psi_1(x)} I(x, y)\, dy$$

where $$I(x, y) = \int_{\phi(x,y)}^{\Psi(x,y)} f(x, y, z)\, dz, \text{ then}$$

$$\iiint_E f(x, y, z)\, dx\, dy\, dz = \int_a^b dx \int_{\phi_1(x)}^{\Psi_1(x)} dy \int_{\phi(x,y)}^{\Psi(x,y)} f(x, y, z)\, dz \qquad \text{..(4)}$$

It is this final formula that reduces the triple integral to an iterated integral.

The variables x, y, z can be interchanged if the corresponding conditions hold.

2. When deriving formula (3) we have taken the domain to be regular with respect to z-axis. If the domain is of a more complicated form, we break up the domain into parts such that each of them is regular with respect to some axis.

A Rule for Limits of Integration. We follow the following steps to reduce a triple integral to an iterated one.

1. Break the domain, if necessary, into sub-domains such that a line parallel to z-axis has at the most two common points with its boundary. In what follows we mention only one such sub-domain.
2. Fix arbitrary x and y and let a line parallel to z-axis cut the boundary of the given domain E at two points with z coordinates $\phi(x, y)$ and $\Psi(x, y)$. The expressions $\phi(x, y)$ and $\Psi(x, y)$ should be taken as the limits of integration with respect to z.
3. The domain of definition of the function of x, y (obtained after integration with respect to z) is now D, the projection of the given domain E on the xy-plane. Let a line parallel to the y-axis cut the bounding curve of D at two points with y coordinates $\phi_1(x)$ and $\Psi_1(x)$. These expressions, $\phi_1(x)$ and $\Psi_1(x)$ form the limits of integration with respect to y.
4. The limits of integral with respect to x can be easily determined.

For example, the triple integral of a function f taken over the sphere $x^2 + y^2 + z^2 = a^2$ is of the form

$$\iiint_E f(x, y, z)\, dx\, dy\, dz = \int_{-a}^{a} dx \int_{-\sqrt{(a^2-x^2)}}^{\sqrt{(a^2-x^2)}} dy \int_{-\sqrt{(a^2-x^2-y^2)}}^{\sqrt{(a^2-x^2-y^2)}} f(x, y, z)\, dz$$

7.10 Change of Variables in Triple Integrals

We have already dealt with the method of changing the variables for the double integrals (§ 5 Ch. 17). The theory underlying the change of variables in triple integrals is exactly the same except that it is more laborious. Here we shall simply indicate the method to be adopted in practical problems.

Let the functions

$$x = X(u, v, w), \quad y = Y(u, v, w), \quad z = Z(u, v, w)$$

map in one-to-one manner, a domain E in cartesian coordinates x, y, z onto a domain E' in the new coordinates u, v, w.

Let $f(x, y, z) = f(X(u, v, w), Y(u, v, w), Z(u, v, w)) = F(u, v, w)$

Then

$$\iiint_E f(x, y, z)\, dx\, dy\, dz = \iiint_{E'} F(u, v, w)\, |J|\, du\, dv\, dw$$

where the Jacobian $J = \dfrac{\partial(x, y, z)}{\partial(u, v, w)}$.

The following two transformations, because of their frequent occurrence, deserve special mention.

(*i*) **Cylindrical polar coordinates.**

$$x = r\cos\theta, \quad y = r\sin\theta, \quad z = z$$

Here, the Jacobian $J = r$.

(*ii*) **Spherical polar coordinates.**

$$x = r \sin\theta \cos\phi, \quad y = r \sin\theta \sin\phi, \quad z = r\cos\theta$$

Jacobian $J = r^2 \sin\theta$.

8. GAUSS'S THEOREM (or *Divergence theorem*) (*Second generalisation of Green's theorem*)

If a three-dimensional regular (or piecewise regular) domain E is bounded by a smooth (or piecewise smooth) oriented surface S, and f, g, h are three functions which along with their partial derivatives f_x, g_y, h_z *are continuous at each point of E and S, then*

$$\iiint_E (f_x + g_y + h_z)\, dx\, dy\, dz = \iint_S (f\, dy\, dz + g\, dz\, dx + h\, dx\, dy)$$

where the surface integral is taken over the exterior of S.

Let us first consider the domain E (Figs. (*i*) and (*ii*) of § 7.8) to be regular with respect to z-axis, and bounded, above and below, by smooth oriented surfaces S_1, S_2 determined by the equations

$$Z = \phi(x, y), \text{ and } Z = \Psi(x, y), \ (\Psi(x, y) \geq \phi(x, y)) \qquad ...(1)$$

and by a lateral cylindrical surface S_3 (which may reduce to the common curve of S_1 and S_2 as in Fig. (*ii*)), with generators parallel to the z-axis. The union of S_1, S_2, S_3 forms the surface S. Let D be the projection of S (or the common projection of S_1 and S_2) on the xy-plane.

As we are considering the outer side of the surface S, the outward drawn normals of S_1 and S_2 are in the opposite directions and so S_1 and S_2 are of opposite orientation. Accordingly, if D is the region on the xy-plane on which S_1 projects, then $(-D)$ is the region on which S_2 projects.

Thus, we have

$$\iiint_E h_z\, dx\, dy\, dz = \iint_D \left[\int_{\phi(x,y)}^{\Psi(x,y)} h_z dz \right] dx\, dy$$

$$= \iint_D [h(x, y, \Psi) - h(x, y, \phi)]\, dx\, dy$$

$$= \iint_D h(x, y, \Psi)\, dx\, dy - \iint_D h(x, y, \phi)\, dx\, dy$$

$$= \iint_D h(x, y, \Psi)\, dx\, dy + \iint_{-D} h(x, y, \phi)\, dx\, dy$$

$$= \iint_{S_1} h(x, y, z)\, dx\, dy + \iint_{S_2} h(x, y, z)\, dx\, dy$$

$$= \iint_{S_1} h\, dx\, dy + \iint_{S_2} h\, dx\, dy + \iint_{S_3} h\, dx\, dy$$

where the last surface integral (which is obviously equal to zero) is taken over the outer side of the lateral surface S_3. Thus

$$\iiint_E h_z \, dx \, dy \, dz = \iint_S h \, dx \, dy \quad \text{...(2)}$$

We claim that the relation (2) is also valid for any domain which is piecewise regular with respect to z-axis. For, if the domain E is divided, by surfaces into sub-domains each of which is regular with respect to z-axis, relation (2) holds for each sub-domain. When all these relations are added, we get on the left hand side the volume integral over E, and on the right hand side, the surface integral over S (the surface integrals over the dividing surface being taken twice, once over each side, cancel each other).

Since E is regular with respect to all the coordinate axes, considering the domain regular with respect to x-axis and y-axis in turn, we deduce that

$$\iiint_E f_x \, dx \, dy \, dz = \iint_S f \, dy \, dz \quad \text{...(3)}$$

and

$$\iiint_E g_y \, dx \, dy \, dz = \iint_S g \, dz \, dx \quad \text{...(4)}$$

Adding equations (2), (3) and (4), we get

$$\iiint_E (f_x + g_y + h_z) \, dx \, dy \, dz = \iint_S (f \, dy \, dz + g \, dz \, dx + h \, dx \, dy) \quad \text{...(5)}$$

which holds for any regular or piecewise regular domain with a smooth boundary S.

Relation (5) may be expressed in the form

$$\iiint_E (f_x + g_y + h_z) \, dx \, dy \, dz = \iint_S (f \cos\alpha + g \cos\beta + h \cos\gamma) \, dS \quad \text{...(6)}$$

where $\cos\alpha, \cos\beta, \cos\gamma$ are the direction cosines of the outward drawn unit normal of S.

Note: (*Vectorial formulation*). Let $\mathbf{F} = \mathbf{i}P + \mathbf{j}Q + \mathbf{k}R$ be a vector function defined on E and S, then

$$\operatorname{div} \mathbf{F} = \frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z}$$

Let $\mathbf{n} = \mathbf{i} \cos\alpha + \mathbf{j} \cos\beta + \mathbf{k} \cos\gamma$

be an outward drawn unit normal of the surface S, then

$$\mathbf{F} \cdot \mathbf{n} = P \cos\alpha + Q \cos\beta + R \cos\gamma$$

Hence, Gauss's theorem may be expressed as

$$\iiint_E \operatorname{div} \mathbf{F} \, dV = \iint_S \mathbf{F} \cdot \mathbf{n} \, dS \quad \text{...(7)}$$

where dV denotes an element of volume.

Relation (7) shows that Gauss's theorem may be stated as:

"The integral of the divergence of a vector function $\mathbf{F}$ *over some volume is equal to the vector flux through the surface bounding the given volume."*

8.1 Applications of Gauss's Theorem

I. Evaluation of surface integrals. We know that a surface integral can generally be evaluated by reducing it to the corresponding double integral. But there are cases where this is inconvenient. In such cases, it is generally advisable to reduce a surface integral over a closed surface to a triple integral by means of Gauss's Theorem.

ILLUSTRATIONS

1. Evaluate the surface integral

$$I = \iint_S (x^3\, dy\, dz + y^3\, dz\, dx + z^3\, dx\, dy)$$

over the sphere $x^2 + y^2 + z^2 = a^2$.

By Gauss's Theorem, we have

$$I = 3\iiint_E (x^2 + y^2 + z^2)\, dx\, dy\, dz$$

where the domain E is the sphere $x^2 + y^2 + z^2 \le a^2$.

Changing to spherical polars, we get

$$I = 3\int_0^{2\pi} d\phi \int_0^{\pi} d\theta \int_0^{a} r^4 \sin\theta\, dr = \frac{12}{5}\pi a^5$$

2. Evaluate the surface integral

$$I = \iint_S z\, dy\, dz + x\, dz\, dx + y\, dx\, dy$$

over a closed surface S.

By Gauss's Theorem, the integral reduces to the triple integral (over the domain bounded by the surface S) whose integrand is identically equal to zero. Hence $I = 0$ for any closed surface S.

II. Volume of a solid by a surface integral. Just as Green's Theorem enables us to express the area of a plane figure as a line integral along its boundary, the Gauss' Theorem helps us to find an expression of the volume of a solid in the form of a surface integral over the closed surface bounding the solid.

Let us choose three functions f, g, h so that

$$f_x + g_y + h_z = 1$$

Then we obtain,

$$\iint_S f\ y\ z + g\ dz\ dy\ dx + h\ dx\ dy = \iiint_E dx\ dy\ dz = V$$

where V is the volume of the domain E bounded by S, and the surface integral is taken over the outer side of the surface S.

In particular, putting

$$f(x, y, z) = x,\ g = 0,\ h = 0$$

we get

$$V = \iint_S x\, dy\, dz \qquad \text{...(8)}$$

Similarly

$$V = \iint_S y \, dz \, dx, \quad V = \iint_S z \, dx \, dy \qquad ...(9)$$

Adding these results, or putting

$$f(x, y, z) = \frac{1}{3}x, \; g(x, y, z) = \frac{1}{3}y, \; h(x, y, z) = \frac{1}{3}z$$

we get

$$V = \frac{1}{3}\iint_S (x \, dy \, dz + y \, dz \, dx + z \, dx \, dy) \qquad ...(10)$$

the integration being taken over the outer side of S in each case.

ILLUSTRATION

The volume V of the sphere $x^2 + y^2 + z^2 = a^2$ is given by

$$V = \frac{1}{3}\iint_S (x \, dy \, dz - y \, dz \, dx + z \, dx \, dy)$$

$$= 2\iint_D \sqrt{a^2 - x^2 - y^2} \, dx \, dy$$

where D is the circle $x^2 + y^2 \le a^2$

$$= 2\iint_0^{2\pi} d\theta \int_0^a (\sqrt{a^2 - r^2}) \, r \, dr = \frac{4}{3}\pi a^2$$

Example 22. Compute the integral

$$\iiint_E xyz \, dx \, dy \, dz$$

over a domain bounded by $x = 0, y = 0, z = 0, x + y + z = 1$.

■ The domain is regular and bounded, above and below by $z = 1 - x - y$ and $z = 0$. Its projection D on the xy-plane is a triangle bounded by $x = 0, y = 0, y = 1 - x$.

$$\therefore \qquad \iiint_E xyz \, dx \, dy \, dz = \iint_D \left[\int_0^{1-x-y} xyz \, dz\right] dx \, dy$$

$$= \int_0^1 dx \int_0^{1-x} \frac{1}{2}xy\,(1 - x - y)^2 \, dy$$

$$= \int_0^1 \frac{x}{24}(1 - x)^4 \, dx = \frac{1}{720}$$

Example 23. Evaluate

$$I = \iint_S (x \cos\alpha + y \cos\beta + z^2 \cos\gamma) \, dS,$$

where S denotes the closed surface bounded by the cone $x^2 + y^2 = z^2$ and the plane $z = 1$; and $\cos\alpha, \cos\beta, \cos\gamma$ are direction cosines of the outward drawn normal of S.

■ By Gauss's Theorem

$$I = \iiint_E (2 + 2z)\, dx\, dy\, dz$$

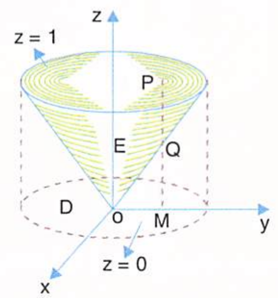


Fig. 13

where E is the domain bounded by $x^2 + y^2 = z^2$ and $z = 1$

$$= \iint_D dx\, dy \int_{\sqrt{(x^2+y^2)}}^{1} (2 + 2z)\, dz$$

$$= \iint_D [3 - 2\sqrt{x^2 + y^2} - (x^2 + y^2)]\, dx\, dy$$

$$= \int_0^{2\pi} d\theta \int_0^1 (3 - 2r - r^2)\, r\, dr = \frac{7\pi}{6}$$

Note: The given surface integral is same as

$$\iint_S (x\, dy\, dz + y\, dz\, dx + z^2\, dx\, dy)$$

Example 24. Evaluate the surface integral

$$I = \iint_S (y^2 z\, dx\, dy + xz\, dy\, dz + x^2 y\, dz\, dx),$$

where S is the outer side of the surface situated in the first octant and formed by the paraboloid of revolution $z = x^2 + y^2$, cylinder $x^2 + y^2 = 1$ and the coordinate planes.

■ Using Gauss's Theorem

$$I = \iiint_V (x^2 + y^2 + z)\, dx\, dy\, dz$$

where V is the volume enclosed by S.

$$= \iint_D \left[\int_0^{x^2+y^2} (x^2 + y^2 + z)\, dz\right] dx\, dy$$

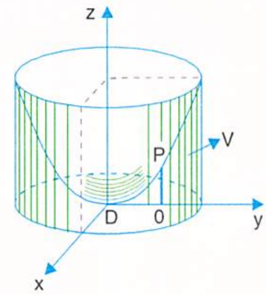


Fig. 14

where $D(x^2 + y^2 \leq 1)$ is the projection on the xy-plane of the domain enclosed by S.

$$= \frac{3}{2} \iint_D (x^2 + y^2)^2 \, dx \, dy$$

$$= \frac{3}{2} \int_0^{\pi/2} d\theta \int_0^1 r^4 r \, dr = \frac{\pi}{8}$$

Example 25. Compute the volume of the ellipsoid

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$$

■ The ellipsoid is bounded above and below by the surfaces

$$z = c\sqrt{1 - \frac{x^2}{a^2} - \frac{y^2}{b^2}}, \text{ and } z = -c\sqrt{1 - \frac{x^2}{a^2} - \frac{y^2}{b^2}}$$

and its projection on the xy-plane is the ellipse $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$

$$\therefore \qquad \text{Volume} = \iiint_E dx \, dy \, dz$$

$$= \iint_D dx \, dy \int_{-c\sqrt{(1-x^2/a^2-y^2/b^2)}}^{c\sqrt{(1-x^2/a^2-y^2/b^2)}} dz$$

$$= 2c \iint_D \sqrt{1 - \frac{x^2}{a^2} - \frac{y^2}{b^2}} \, dx \, dy$$

Putting $\qquad x/a = r \cos\theta, \; y/b = r \sin\theta,$

$$J = abr$$

$$\therefore \qquad \text{Volume} = 2\, abc \int_0^{2\pi} d\theta \int_0^1 (\sqrt{1 - x^2}) \, r \, dr$$

$$= 2abc \cdot 2\pi \left[-\frac{1}{3} (1 - r^2)^{3/2} \right]_0^1 = \frac{4}{3} \pi abc$$

Example 26. Compute the volume of the solid bounded by the sphere $x^2 + y^2 + z^2 = 4$ and the surface of the paraboloid $x^2 + y^2 = 3z$.

■ The two surfaces intersect at $z = 1$.

The domain E, under consideration, is bounded, above and by the two surfaces $z = \sqrt{4 - x^2 - y^2}$, and $z = \frac{1}{3}\left(x^2 + y^2\right)$, and its projection D on the xy-plane is the circle.

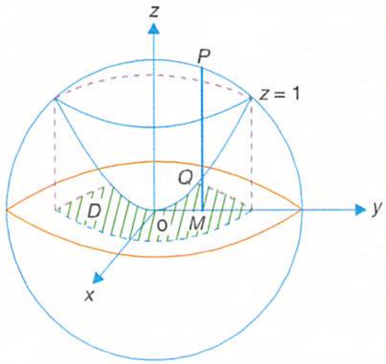


Fig. 15

$$x^2 + y^2 \leq 3$$

$$\therefore \qquad \text{Volume} = \iiint_E dx\, dy\, dz$$

$$= \iint_D dx\, dy \int_{\frac{1}{3}\sqrt{(x^2+y^2)}}^{\sqrt{(4-x^2-y^2)}} dz$$

$$= \iint_D \left[\sqrt{4 - x^2 - y^2} - \frac{1}{3}(x^2 + y^2)\right] dx\, dy$$

Changing to polars, (D being the circle, $x^2 + y^2 \leq 3$),

$$= \int_0^{2\pi} d\theta \int_0^{\sqrt{3}} \left(\sqrt{4 - r^2} - \frac{r^2}{3}\right) r\, dr = \frac{19}{6}\pi$$

.

EXERCISE

1. Compute the integrals

(i) $\int_0^a dx \int_0^x dy \int_0^y xyz\, dz$ (ii) $\int_0^a dx \int_0^x dy \int_0^{xy} x^3y^3z\, dz$

2. Compute $\iiint_E xy\, dx\, dy\, dz$, when E is the domain bounded by $z = xy$, $x + y = 1$, $z = 0$ $(z \geq 0)$.

3. Evaluate $\iiint_E y\cos(z+x)\,dx\,dy\,dz$, E being the domain bounded by the cylinder $y=\sqrt{x}$ and the planes, $y=0$, $z=0$, $x+z=\pi/2$.

4. Compute the integral $\iiint \dfrac{dx\,dy\,dz}{(x+y+z+1)^3}$, over the volume bounded by the planes, $x=0$, $y=0$, $z=0$, $x+y+z=1$.

5. Find the volume of a circular cone with height h and radius a.

6. Evaluate the volume of enclosed region:
   (*i*) Cone $x^2+y^2=z^2$ and the plane $z=1$,
   (*ii*) Parabolic cylinder $z=4-x^2$ and the planes, $x=0$, $y=0$, $z=0$, $y=6$.

7. Calculate the volume of the solid bounded by a surface with the equation $(x^2+y^2+z^2)^2=a^3x$.

8. Find the volume of the solid bounded by
   (*i*) The paraboloid $z=x^2+y^2$, cylinder $y=x^2$ and the planes $y=1$, $z=0$.
   (*ii*) The cylinders $z=4-y^2$, $y=\frac{1}{2}x^2$, and the plane $z=0$.
   (*iii*) The cylinders $x^2+y^2=a^2$, $z=x^3/b^2$, and the plane $z=0$, $(x\geq 0)$.
   (*iv*) The cylinder $x^2+y^2=2ax$ and the paraboloid $y^2+z^2=4ax$.

Use Gauss's Theorem to evaluate the integrals.

9. $\iint_S (x\,dy\,dz+y\,dz\,dx+z\,dx\,dy)$, S being the outer side of the cube $[0, a; 0, a; 0, a]$.

10. $\iint_S [(x^3-yz)\cos\alpha-2x^2y\cos\beta+2\cos\gamma]\,dS$; taken over the outer surface of the cube bounded by the planes $x=0, x=a; y=0, y=a; z=0, z=a$. $\cos\alpha$, $\cos\beta$, $\cos\gamma$ are the direction cosines of the outward drawn normal.

11. $\iint_S (ax\cos\alpha+by\cos\beta+cz\cos\gamma)\,dS$ where S is the surface of the sphere $x^2+y^2+z^2=1$, and α, β, γ have their usual meaning.

12. $\iint_S (y^2z^2\,dy\,dz+z^2x^2\,dz\,dx+x^2y^2\,dx\,dy)$, S being the surface of the sphere $x^2+y^2+z^2=1$ above the xy-plane.

13. $\iint_S (x^2\,dy\,dz+y^2\,dz\,dx+z^2\,dx\,dy)=0$, taken over the surface of the ellipsoid
$$x^2/a^2+y^2/b^2+z^2/c^2=1.$$

14. $\iint_S (x^3\cos\alpha+y^3\cos\beta+z^3\cos\gamma)\,dS$, where S is the surface of a sphere of radius a with centre at the origin and α, β, γ are the angles of inclination to the coordinate axes of the normal to the surface.

15. $\iint (xz\,dx\,dy+xy\,dy\,dz+yz\,dz\,dx)$, taken over the pyramid formed by the planes, $x=0$, $y=0$, $z=0$, $x+y+z=1$.

16. $\iint_S (xz\,dy\,dz+xy\,dz\,dx+yz\,dx\,dy)$, where S is the outer surface situated in the first octant and formed by the cylinder $x^2+y^2=a^2$ and the planes, $x=0$, $y=0$, $z=0$, $z=h$.

ANSWERS

1. (i) $a^6/48$ (ii) $a^{11}/110$
2. $1/180$
3. $(\pi^2 - 8)/16$
4. $\frac{1}{2}\log 2 - 5/16$
5. $\pi a^2 h/3$
6. $\pi/3$
7. $\pi a^3/3$
8. (i) $88/105$ (ii) $256/21$ (iii) $4a^5/15b^2$ (iv) $2a^3(3\pi + 8)/3$
9. $3a^3$
10. $\frac{1}{3}a^5$
11. $4\pi(a + b + c)/3$
12. $\pi/24$
14. $12\pi a^5/5$
15. $1/8$
16. $a^2h(2a/3 + \pi h/8)$.

Example 27. Compute

$$I = \iiint \sqrt{1 - \frac{x^2}{a^2} - \frac{y^2}{b^2} - \frac{z^2}{c^2}}\, dx\, dy\, dz$$

taken over the region $x^2/a^2 + y^2/b^2 + z^2/c^2 \le 1$.

■ Let us change to spherical polar coordinates, where

$$x/a = r \sin\theta \cos\phi,\ y/b = r \sin\theta \sin\phi,\ z/c = r\cos\theta$$
$$0 \le r \le 1,\ \ 0 \le \theta \le \pi,\ \ 0 \le \phi \le 2\pi$$

so that

$$J = \frac{\partial(x, y, z)}{\partial(r, \theta, \phi)} = abcr^2 \sin\theta$$

$$\therefore \qquad I = abc \int_0^{2\pi} d\phi \int_0^{\pi} \sin\theta\, d\theta \int_0^1 (\sqrt{1 - r^2})r^2\, dr = \frac{\pi^2 abc}{4}.$$

Example 28. Show that

$$I = \iiint x^{l-1} y^{m-1} z^{n-1} (1 - x - y - z)^{p-1}\, dx\, dy\, dz,\ \ (l, m, n, p \ge 1)$$

taken over the tetrahedron bounded by the planes, $x = 0$, $y = 0$, $z = 0$, $x + y + z = 1$ is $\dfrac{\Gamma(l)\,\Gamma(m)\,\Gamma(n)\,\Gamma(p)}{\Gamma(l + m + n + p)}$.

■ The given integral is same as

$$\int_0^1 dx \int_0^{1-x} dy \int_0^{1-x-y} x^{l-1} y^{m-1} z^{n-1} (1 - x - y - z)^{p-1}\, dz.$$

First method. Let us put $x + y + z = u$, $x + y = uv$, $x = uvw$, *i.e.*, $x = uvw$, $y = uv\ (1 - w)$, $z = u(1 - v)$

It may be seen that when x, y, z are positive and $x + y + z \le 1$, then each of u, v, w lie between 0 and 1, and conversely. So the given region is fully described when $0 \le u \le 1, 0 \le v \le 1, 0 \le w \le 1$.

Also, then

$$|J| = \left|\frac{\partial(x, y, z)}{\partial(u, v, w)}\right| = |-u^2 v| = u^2 v$$

$$\therefore \qquad I = \int_0^1 u^{l+m+n-1}(1-u)^{p-1}\,du \int_0^1 v^{l+m-1}(1-v)^{n-1}\,dv$$

$$\int_0^1 x^{l-1}(1-w)^{m-1}\,dw$$

$$= \beta(l+m+n, p)\cdot\beta(l+m, n)\cdot\beta(l, m) = \frac{\Gamma(l)\,\Gamma(m)\,\Gamma(n)\,\Gamma(p)}{\Gamma(l+m+n+p)}.$$

Second method. Put $x = u$, $y = (1 - u)v$, $z = (1 - u)(1 - v)w$ so that

$1 - x - y - z = (1 - u)(1 - v)\ (1 - w)$ and

$$|J| = \left|\frac{\partial(x, y, z)}{\partial(u, v, w)}\right| = (1-u)^2(1-v)$$

$$\therefore \qquad I = \int_0^1 u^{l-1}(1-u)^{m+n+p-1}\,du \int_0^1 v^{m-1}(1-v)^{n+p-1}\,dv$$

$$\int_0^1 w^{n-1}(1-w)^{p-1}\,dw$$

$$= \beta(l, m+n+p)\cdot\beta(m, n+p)\cdot\beta(n, p) = \frac{\Gamma(l)\,\Gamma(m)\,\Gamma(n)\,\Gamma(p)}{\Gamma(l+m+n+p)}$$

Note: The usefulness of the two sets of substitution lies in the fact that the new integrals are with constant limits (ref. solved example 26, Ch.17)

Example 29. Evaluate

$$\iiint_E z^2\,dx\,dy\,dz$$

taken over the region common to the surfaces

$$x^2 + y^2 + z^2 = a^2, \text{ and } x^2 + y^2 = ax$$

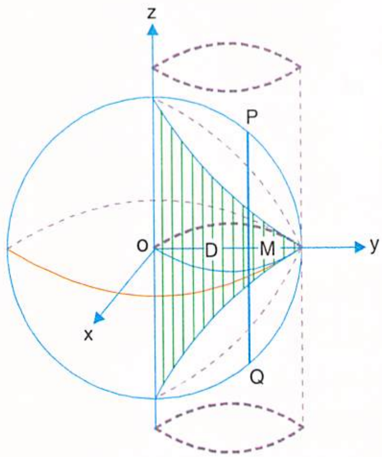
■ The region is bounded, above and below by the surfaces

$$z = \sqrt{a^2 - x^2 - y^2} \text{ and } z = -\sqrt{a^2 - x^2 - y^2}$$

and its projection on the xy-plane is the circular domain $D \equiv x^2 + y^2 \le ax$.

$$\therefore \qquad \iiint_E z^2\,dx\,dy\,dz = \iint_D dx\,dy \int_{-\sqrt{(a^2-x^2-y^2)}}^{\sqrt{(a^2-x^2-y^2)}} z^2\,dz$$

$$= \frac{2}{3} \iint_D (a^2 - x^2 - y^2)^{3/2} \, dx \, dy$$



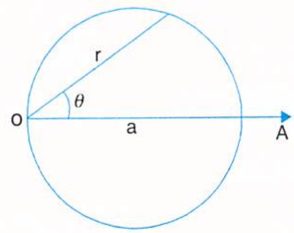


Fig. 16

Changing to polars, the region D becomes the circle,

$$r = a \cos \theta$$

where, $\quad 0 \le \theta \le \pi, 0 \le r \le a \cos \theta$

$$= \frac{2}{3} \int_0^{\pi} d\theta \int_0^{a \cos \theta} (a^2 - r^2)^{3/2} r \, dr$$

$$= \frac{2}{15} a^5 \int_0^{\pi/2} (1 - \sin^5 \theta) \, d\theta = \frac{2a^5(15\pi - 16)}{225}.$$

Example 30. Evaluate

$$I = \iiint_E (y^2 z^2 + z^2 x^2 + x^2 y^2) \, dx \, dy \, dz$$

taken over the domain bounded by the cylinder $x^2 + y^2 = 2ax$, and the cone $z^2 = k^2(x^2 + y^2)$.

■ The domain E is bounded above and below by the surface

$$z = k\sqrt{x^2 + y^2}$$

and $\quad z = -k\sqrt{x^2 + y^2}$

and its projection on the xy-plane is the circular domain D, $x^2 + y^2 \le 2ax$.

$$\therefore \quad I = \iint_D dx \, dy \times \int_{-k\sqrt{(x^2+y^2)}}^{k\sqrt{(x^2+y^2)}} \{(x^2 + y^2) z^2 + x^2 y^2\} \, dz$$

$$= 2 \iint_D [\tfrac{1}{3}(x^2 + y^2)^2 k^2 + x^2 y^2] k\sqrt{x^2 + y^2} \, dx \, dy.$$

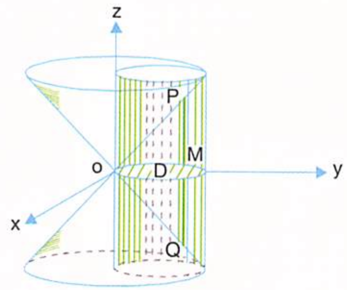


Fig. 17

Changing to polars,

$$= \frac{4k}{3} \int_0^{\pi/2} d\theta \int_0^{2a\cos\theta} (k^2 + 3\cos^2\theta \sin^2\theta)\, r^6\, dr$$

$$= \frac{512}{21} ka^7 \int_0^{\pi/2} (k^2 \cos^7\theta + 3\cos^9\theta \sin^2\theta)\, d\theta$$

$$= \frac{8192}{735} ka^7 \left(k^2 + \frac{8}{33}\right)$$

Example 31. Integrate $1/xyz$ throughout the volume enclosed by the six spheres, $x^2 + y^2 + z^2 = ax$, $a'x$; by, $b'y$; cz, $c'z$; a, a'; b, b'; c, c' being all positive.

■ Let

$$\frac{(x^2 + y^2 + z^2)}{x} = u, \frac{(x^2 + y^2 + z^2)}{y} = v, \frac{(x^2 + y^2 + z^2)}{z} = w$$

so that

$$x = \frac{1}{u\sum u^{-2}}, y = \frac{1}{v\sum u^{-2}}, z = \frac{1}{w\sum u^{-2}}.$$

If we take u, v, w as the new coordinate system, the new domain of integration is the parallelepiped, $[a, a'; b, b'; c, c']$.

The Jacobian,

$$|J| = \begin{vmatrix} \dfrac{2u^{-2} - \sum u^{-2}}{u^2(\sum u^{-2})^2} & \dfrac{2v^{-2}}{uv(\sum u^{-2})^2} & \dfrac{2w^{-2}}{uw(\sum u^{-2})^2} \\ \dfrac{2u^{-2}}{uv(\sum u^{-2})^2} & \dfrac{2v^{-2} - \sum u^{-2}}{v^2(\sum u^{-2})^2} & \dfrac{2w^{-2}}{vw(\sum u^{-2})^2} \\ \dfrac{2u^{-2}}{uw(\sum u^{-2})^2} & \dfrac{2v^{-2}}{vw(\sum u^{-2})^2} & \dfrac{2w^{-2} - \sum u^{-2}}{w^2(\sum u^{-2})^2} \end{vmatrix}$$

$$= \frac{1}{u^2 v^2 w^2 (\sum u^{-2})^3}$$

$$\therefore \qquad \iiint_E \frac{1}{xyz}\, dx\, dy\, dz = \int_a^{a'} \frac{1}{u}\, du \int_b^{b'} \frac{1}{v}\, dv \int_c^{c'} \frac{1}{w}\, dw$$

$$= \log \frac{a'}{a} \log \frac{b'}{b} \log \frac{c'}{c}\,.$$

Example 32. Evaluate $\iint \mathbf{F} \cdot \mathbf{n}\, dS$, where S is the entire surface $x^2 + y^2 = 1$, $z = 0$, $z = x + 2$, and $\mathbf{n}$ is the outward drawn unit normal, and

$$\mathbf{F} = 2x\mathbf{i} - 3y\mathbf{j} + z\mathbf{k}.$$

- By Gauss's Divergence Theorem
  (This is solved as surface integral in Ex. 14).

$$\iint_S \mathbf{F} \cdot \mathbf{n}\, dS = \iiint_V \text{div } \mathbf{F}\, dv$$

$$= \iiint \left(\frac{\partial}{\partial x}(2x) + \frac{\partial}{\partial y}(-3y) + \frac{\partial}{\partial z}(z) \right) dx\, dy\, dz = 0.$$

This shows that not only Gauss' Theorem is very useful in changing volume integral to surface integral but also simplifies the problem by changing surface integral to volume integral.

Example 33. (Same as example 11). Evaluate the surface integral

$$\iint z \cos \gamma\, dS,$$

over the outer side of the sphere $x^2 + y^2 + z^2 = 1$, where γ is the inclination of the normal at any point of the sphere with z-axis.

- Since $\cos \gamma = \mathbf{k} \cdot \mathbf{n}$, where $\mathbf{n}$ is the outward drawn unit normal to the surface.
  $\therefore$ By Gauss' Divergence theorem, we have

$$\iint_S z \cos \gamma\, dS = \iint_S z\, \mathbf{k} \cdot \mathbf{n}\, dS = \iiint_V \text{div } (z\, \mathbf{k})\, dx\, dy\, dz$$

$$= \iiint_V 1 \cdot dx\, dy\, dz$$

$$= \frac{4}{3} \pi (1)^3 \text{ (volume of the sphere with unit radius)}$$

$$= 4\pi/3\,.$$

Example 34. Evaluate $\int (x^2 + y^2)\, ds$ and $\int (x^2 + y^2)\, dS$, where S is the surface of the sphere $x^2 + y^2 + z^2 = a^2$.

- Taking, $\mathbf{n} = \left(\frac{x}{a}, \frac{y}{a}, \frac{z}{a} \right)$, and $\mathbf{F} = (ax, ay, 0)$, we have

CHAPTER 18

$$\int (x^2 + y^2)\, dS = \iint_S \mathbf{F} \cdot \mathbf{n}\, dS$$

$$= \iiint_V \text{div } \mathbf{F}\, dV, \text{ by Gauss' theorem}$$

$$= \iiint_V (a + a)\, dx\, dy\, dz = 2a\, \frac{4}{3}\, \pi a^3 = \frac{8}{3}\, \pi a^4,$$

Now

$$\int_S (x^2 + y^2)\, dS = \int_S (x^2 + y^2)\, \mathbf{n}\, ds$$

$$= \iint_S (x^2 + y^2)\, (\mathbf{i}\, dy\, dz + \mathbf{j}\, dz\, dx + \mathbf{k}\, dx\, dy)$$

Taking the parametric co-ordinates,

$x = a \sin\theta \cos\phi,\ y = a \sin\theta \sin\phi,\ z = a \cos\theta,$

where $0 \le \theta \le \pi,\ 0 \le \phi \le 2\pi$, on the sphere $x^2 + y^2 + z^2 = a^2$.

$$\therefore \int_S (x^2 + y^2)\, dS = \int_0^{\pi} \int_0^{2\pi} a^2 \sin^2\theta \left[\mathbf{i}\, \frac{\partial(y, z)}{\partial(\theta, \phi)} + \mathbf{j}\, \frac{\partial(z, x)}{\partial(\theta, \phi)} + \mathbf{k}\, \frac{\partial(x, y)}{\partial(\theta, \phi)} \right] d\theta\, d\phi$$

$$= \int_0^{\pi} \int_0^{2\pi} a^2 \sin^2\theta\, (\mathbf{i}\, a^2 \sin^2\theta \cos\phi + \mathbf{j}\, a^2 \sin^2\theta \sin\phi + \mathbf{k}\, a^2 \sin\theta \cos\theta)\, d\theta\, d\phi$$

$$= 0.$$

EXERCISE

1. Evaluate the integrals by passing over to cylindrical or spherical polar coordinates.

 (i) $\int_0^1 dx \int_{-\sqrt{(1-x^2)}}^{\sqrt{(1-x^2)}} dy \int_0^a dz$

 (ii) $\int_0^2 dx \int_0^{\sqrt{(2x-x^2)}} dy \int_0^a z\sqrt{x^2 + y^2}\, dz$

 (iii) $\int_{-a}^{a} dx \int_{-\sqrt{(a^2-x^2)}}^{\sqrt{(a^2-x^2)}} dy \int_0^{\sqrt{(a^2-x^2-y^2)}} (x^2 + y^2)\, dz$

 (iv) $\int_0^1 dx \int_0^{\sqrt{(1-x^2)}} dy \int_0^{\sqrt{(1-x^2-y^2)}} \sqrt{x^2 + y^2 + z^2}\, dz$

2. Compute $\iiint_E (x^2 + y^2)\, dx\, dy\, dz$, where E is specified by $z \ge 0,\ a^2 \le x^2 + y^2 + z^2 \le b^2$.

3. Compute $\iiint_E \dfrac{dx\, dy\, dz}{\sqrt{x^2 + y^2 + (z-2)^2}}$, where E is the sphere $x^2 + y^2 + z^2 \le 1$.

4. Show that

$$\iiint_E (ax + by + cz)^2 \, dx\, dy\, dz = \frac{4}{15}\pi (a^2 + b^2 + c^2)$$

where domain E is the sphere $x^2 + y^2 + z^2 \le 1$.

5. Show that

$$\iiint (lx^2 + my^2 + nz^2)\, dx\, dy\, dz$$

taken throughout the sphere $x^2 + y^2 + z^2 = a^2$, is $4\pi(l + m + n)a^5/15$.

6. Prove that

$$\iiint_E z\, dx\, dy\, dz = \frac{\pi}{4} h^4 \cot\alpha \cot\beta,$$

where E is the domain bounded by the cone $z^2 = x^2 \tan^2\alpha + y^2 \tan^2\beta$ and the planes $z = 0$, $z = h$.

Further, if E_1 is the part of the domain E for which x, y and z are positive, show that

$$\iiint_{E_1} xyz\, dx\, dy\, dz = \frac{1}{48} h^6 \cot^2\alpha \cot^2\beta.$$

7. Show that

$$\iiint \frac{dx\, dy\, dz}{\sqrt{1 - x^2 - y^2 - z^2}} = \frac{\pi^2}{8}$$

integral being extended to all the positive values of the variables for which the expression is real.

8. Show that

$$\iiint e^{\sqrt{(x^2/a^2 + y^2/b^2 + z^2/c^2)}}\, dx\, dy\, dz = 4\pi abc(e - 2)$$

integral being taken over the region $x^2/a^2 + y^2/b^2 + z^2/c^2 \le 1$.

9. Prove that

$$\iiint E \sqrt{\frac{1 - x^2 - y^2 - z^2}{1 + x^2 + y^2 + z^2}}\, dx\, dy\, dz = \frac{\pi}{8}\left[\beta\left(\frac{3}{4}, \frac{1}{2}\right) - \beta\left(\frac{5}{4}, \frac{1}{2}\right)\right],$$

where E is the domain for which x, y, z are all positive and $x^2 + y^2 + z^2 \le 1$.

10. Evaluate $\iiint z\, dx\, dy\, dz$, over the region $x^2 + y^2 \le z^2$, $x^2 + y^2 + z^2 \le 1$, $z \ge 0$.

11. Compute the volumes of the solids bounded by

(*i*) The cylinders $z = 4 - y^2$, and $z = y^2 + 2$, and the planes $x = -1$, $x = 2$

(*ii*) The paraboloids $z = x^2 + y^2$, and $z = x^2 + 2y^2$, and the planes $y = x$, $y = 2x$, $x = 1$.

12. (*i*) The paraboloids $(x - 1)^2 + y^2 = z$, and planes $2x + z = 2$.

[*Projection of the solid on the xy-plane is a circle*.]

(ii) The sphere $x^2 + y^2 + z^2 = a^2$, and the cone $z^2 \sin^2 \alpha = (x^2 + y^2) \cos^2 \alpha$, where α is a constant such that $0 \le \alpha \le \pi$.

[In spherical polar coordinates, the surfaces are $r^2 = a^2$ and $\theta = \alpha$]

13. Show that the volume of the region bounded by the hyperboloid $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} - \dfrac{z^2}{c^2} = 1$, its asymptotic cone $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} - \dfrac{z^2}{c^2} = 0$, and the planes, $z = z_1, z = z_2$ $(z_2 > z_1)$ is $\pi ab(z_2 - z_1)$.

14. Find the volume of the solid bounded by the six cylinders, $z^2 = y, z^2 = 2y; x^2 = z, x^2 = 2z; y^2 = x, y^2 = 2x$.

15. Show that the volume enclosed by the paraboloid $x^2 - y^2 = 2az$, the cylinder $(x^2 + y^2)^2 = a^2(x^2 - y^2)$, and the planes $z = 0$ is $a^3/6$.

16. Show that $\iiint_E (x^3 + y^3 + z^3)\, dx\, dy\, dz = \dfrac{32}{5} \pi a^6$, where E is the interior of the sphere

$$x^2 + y^2 + z^2 - 2a(x + y + z) + 2a^2 = 0.$$

17. Show that $\iiint_E xyz\, dx\, dy\, dz = \dfrac{2}{15} (\lambda^{-2} + \mu^{-2} + \nu^{-2})^{-3}$, where E is the volume common to the spheres

$$x^2 + y^2 + z^2 = 2\lambda x,\ x^2 + y^2 + z^2 = 2\mu y,\ x^2 + y^2 + z^2 = 2\nu z.$$

18. Evaluate $\iiint x^l y^m z^n (1 - ax - by - cz)^p\, dx\, dy\, dz$ over the interior of the tetrahedron formed by the planes $x = 0,\ y = 0,\ z = 0,\ ax + by + cz = 1$.

19. Evaluate the triple integral $\iiint x^{l-1} y^{m-1} z^{n-1}\, dx\, dy\, dz$, where $l, m, n \ge 1$ and the variables are all positive such that

(i) $x/a + y/b + z/c \le h$

(ii) $(x/a)^p + (y/b)^q + (z/c)^r \le 1$

[*Hint*: (i) Put $x/a + y/b + z/c = hu,\ x/a + y/b = huv,\ x/a = huvw\ \ 0 \le u \le 1,\ 0 \le v \le 1,\ 0 \le w \le 1$

(ii) Put $(x/a)^p + (y/b)^q + (z/c)^r = u,\ (x/a)^p + (y/b)^q = uv,\ (x/a)^p = uvw$, so that $(x/a)^p = uvw,\ (y/b)^q = uv(1 - w),\ (z/c)^r = u(1 - v)$

$$0 \le u \le 1,\ 0 \le v \le 1,\ 0 \le w \le 1].$$

20. Evaluate $\iiint x^{l-1} y^{m-1} z^{n-1} f[(x/a)^p + (y/b)^q + (z/c)^r]\, dx\, dy\, dz$ over the region in which x, y, z take positive values subject to the condition that $(x/a)^p + (y/b)^q + (z/c)^r \le h$.

[*Hint*: Put $(x/a)^p + (y/b)^q + (z/c)^r = hu,\ (x/a)^p + (y/b)^q = huv,\ (x/a)^p = huvw$ where $0 \le u \le 1,\ 0 \le v \le 1,\ 0 \le w \le 1$]

21. Prove that

$$\iiint_E z^{-3} \sqrt{\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2}}\, dx\, dy\, dz = \frac{\pi}{pq}$$

where E is the smaller region bounded by the ellipsoid $x^2/a^2 + y^2/b^2 + z^2/c^2 = 1$, and the half-cone $z^2 = p^2x^2 + q^2y^2$, $z > 0$.

22. Show that the entire volume enclosed by the solid $(x/a)^{2/3} + (y/b)^{2/3} + (z/c)^{2/3} = 1$ is $4\pi abc/35$.

23. Prove that the volume in the positive octant bounded by the surface, $(z/c)^m = (x/a)^m + (y/b)^m$, and the planes $x=0$, $y=0$, $z=h$ is equal to

$$\frac{abh^3[\Gamma(1/m)]^2}{6mc^2\,\Gamma(2/m)}$$

24. Prove that the volume of a cone which extends from the origin to the surface $x=f(u, v)$, $y=g(u, v)$, $z=h(u, v)$ is given by $\frac{1}{3}\iint \Delta\, du\, dv$, where

$$\Delta = \begin{vmatrix} x & y & z \\ x_u & y_u & z_u \\ x_v & y_v & z_v \end{vmatrix}$$

Hence prove that the volume of one octant of the cone $x^4 + y^4 = z^4 \tan^2 \alpha$ between its vertex and the surface $x^4 + y^4 + z^4 = 1$ is

$$\frac{\Gamma\left(\frac{1}{4}\right)^2}{24\Gamma\left(\frac{1}{2}\right)}\int^{\alpha}\frac{dt}{\sqrt{\cos t}}.$$

[*Hint*: Put $x^2 = \sin\theta\cos\varphi$, $y^2 = \sin\theta\sin\varphi$, $z^2 = \cos\theta$, where $0 \le \theta \le \alpha$, $0 \le \varphi \le \pi/2$.]

ANSWERS

1. (*i*) $\frac{1}{2}\pi a$, (*ii*) $8a^2/9$, (*iii*) $4\pi a^5/15$, (*iv*) $\pi/8$.

2. $4\pi(b^5 - a^5)/15$ 3. $2\pi/3$ 10. $\pi/8$ 11. (*i*) 8 (*ii*) 7/12

12. (*i*) $\frac{1}{2}\pi$ (*ii*) $2\pi a^3(1-\cos\alpha)/3$ 14. 1/7

18. $\dfrac{\Gamma(l+1)\,\Gamma(m+1)\,\Gamma(n+1)\,\Gamma(p+1)}{a^{l+1}b^{m+1}c^{n+1}\;\Gamma(l+m+n+p+4)}$

19. (*i*) $a^l b^m c^n (h)^{\Sigma l}\,\dfrac{\Gamma l\,\Gamma m\,\Gamma n}{\Gamma(\Sigma l + 1)}$, where $\Sigma l = l + m + n$

(*ii*) $\dfrac{a^l b^m c^n}{pqr}\;\dfrac{\Gamma(l/p)\,\Gamma(m/q)\,\Gamma(n/r)}{\Gamma(l/p + m/q + n/r + 1)}$

20. $\dfrac{a^l b^m c^n}{pqr}(h)^{\Sigma(l/p)}\;\dfrac{\Gamma(l/p)\,\Gamma(m/q)\,\Gamma(n/r)}{\Gamma(l/p + m/q + n/r)}\displaystyle\int_0^1 u^{\Sigma(l/p)} f(hu)\,du,$

where $\Sigma(l/p) = l/p + m/q + n/r$.

19 Metric Spaces

Analysis is mainly concerned with processes pertaining to limits. In this chapter we shall study these processes in a general setting. We recall that in the theory of functions of real variables the notion of distance plays a vital role in formulating the definition of convergence, continuity and differentiability. Metric spaces are sets in which there is defined a notion of '*distance between pair of points*' and they provide the general setting in which we study convergence and continuity. The concept of metric spaces was formulated in 1906 by M. Fréchet.

The first section of this chapter is devoted mainly to basic definitions and important examples of metric spaces. We shall study the concepts of open sets, closed sets, convergence, continuity, compactness and connectedness in the later sections. We shall also prove in this chapter a simple result about complete metric spaces, Banach's Fixed point theorem, that has interesting and important applications in classical analysis.

1. DEFINITIONS AND EXAMPLES

Definition. Let X be a non-empty set. A *metric* on X is a real-valued function $d: X \times X \to \mathbf{R}$ which satisfies the following conditions:

(1) $d(x, y) \geq 0, \forall\ x, y \in X,$

(2) $d(x, y) = 0$ if and only if $x = y,\ \forall\ x, y \in X,$

(3) $d(x, y) = d(y, x),\ \forall\ x, y \in X$ (symmetry),

(4) $d(x, y) \leq d(x, z) + d(z, y), \forall\ x, y, z \in X$ (triangle inequality).

A metric d is also called a *distance function*, and the non-negative real number $d(x, y)$ is to be thought of as the distance between x and y.

A *metric space* is a non-empty set X equipped with a metric d on X and is denoted by the pair (X, d) or simply X. Different metrics can be defined on a single non-empty set and this gives rise to distinct metric spaces.

ILLUSTRATIONS

1. The function $d : \mathbf{R} \times \mathbf{R} \to \mathbf{R}$ defined by

$$d(x, y) = |\,x - y\,|,\ \ \forall\ x, y \in \mathbf{R}$$

is a metric on the set $\mathbf{R}$ of all real numbers, since for $x, y, z \in \mathbf{R}$, we have

$$d(x, y) = |\,x - y\,| \leq |\,x - z\,| + |\,z - y\,| = d(x, z) + d(z, y)$$

The number $d(x, y)$ is, of course, the usual '*distance*' between the points x, y on the real line. Therefore d is sometimes referred to as the *usual metric* on **R**.

2. The function d defined by

$$d(z_1, z_2) = |z_1 - z_2|, \quad \forall\, z_1, z_2 \in \mathbf{C}$$

is a metric on the set C of all complex numbers. To prove the triangle inequality it is sufficient to prove that

$$|z_1 + z_2| \le |z_1| + |z_2|, \quad \forall\, z_1, z_2 \in \mathbf{C}$$

This follows from the following:

$$\begin{aligned} |z_1 + z_2|^2 &= (z_1 + z_2)(\bar{z}_1 + \bar{z}_2) \\ &= z_1\bar{z}_1 + z_1\bar{z}_2 + z_2\bar{z}_1 + z_2\bar{z}_2 \\ &= z_1\bar{z}_1 + 2\,Re\,(z_1\,\bar{z}_2) + z_2\bar{z}_2 \\ &\le |z_1|^2 + 2|z_1|\,|z_2| + |z_2|^2 = (|z_1| + |z_2|)^2. \end{aligned}$$

3. Let X be an arbitrary non-empty set. The function d defined by

$$d(x, y) = \begin{cases} 1, & \text{if } x \ne y \\ 0, & \text{if } x = y \end{cases} \quad \forall\, x, y \in X$$

is a metric on X and is called the *discrete* (*trivial*) *metric* on X, and (X, d) is called the *discrete metric space* or the *trivial metric space*.

4. The set l_∞ of all bounded sequences $\{x_n\}$ of real numbers with the function d defined by

$$d(\{x_n\}, \{y_n\}) = \sup\{|x_n - y_n| : n \in N\}, \forall\, \{x_n\}, \{y_n\} \in l_\infty$$

is a metric on l_∞.

For the triangle inequality, we have $\forall\, \{x_n\}, \{y_n\}, \{z_n\} \in l_\infty$

$$|x_n - y_n| = |x_n - z_n + z_n - y_n| \le |x_n - z_n| + |z_n - y_n|,$$

$$\therefore \quad \sup_n |x_n - y_n| \le \sup_n |x_n - z_n| + \sup_n |z_n - y_n|$$

i.e., $d(\{x_n\}, \{y_n\}) \le d(\{x_n\}, \{z_n\}) + d(\{z_n\}, \{y_n\})$.

5. The set $C[0, 1]$ consisting of all real-valued continuous functions defined on $[0, 1]$ with the function d given by

$$d(f, g) = \int_0^1 |f(x) - g(x)|\,dx, \quad \forall\, f, g \in C[0, 1]$$

is a metric space.

6. The set $C\,[0, 1]$ with another metric d defined by

$$d(f, g) = \sup_{x \in [0,1]} |f(x) - g(x)|, \quad \forall\, f, g \in C[0, 1]$$

is a metric space. The metric d is called the *Tehebyshev metric* or *sup metric*.

7. The set **C** of complex numbers with the metric d defined by

$$d(x, y) = \begin{cases} |x| + |y|, & \text{if } x \neq y \\ 0, & \text{if } x = y \end{cases}$$

is a metric space.

In the following examples the conditions (1), (2) and (3) of definition of the metric are easy to verify and are left to the reader. We shall verify only the triangle inequality.

Example 1. Show that the set $\mathbf{R}^n$ of all ordered n-tuples with the function d defined by

$$d(x, y) = \left(\sum_{i=1}^{n} (x_i - y_i)^2\right)^{1/2}, \forall\ x = (x_1, x_2 \ldots, x_n),\ y = (y_1, y_2 \ldots, y_n) \in \mathbf{R}^n$$

is a metric space (d is called the *Euclidean metric* on $\mathbf{R}^n$).

■ To prove the triangle inequality we shall use the following *Cauchy-Schwarz inequality*:

$$\left|\sum_{k=1}^{n} a_k\, b_k\right| \leq \sqrt{\sum_{k=1}^{n} a_k^2 \sum_{k=1}^{n} b_k^2}$$

where $a = (a_1, a_2, \ldots, a_n)$, and $b = (b_1, b_2, \ldots, b_n) \in \mathbf{R}^n$

If $b_k = 0$, for $1 \leq k \leq n$, then there is nothing to prove.

Assume that $b_k \neq 0$ for some k, $1 \leq k \leq n$, then $\sum_{k=1}^{n} b_k^2 > 0$.

If x is any real number, then we have

$$\sum_{k=1}^{n} (a_k - xb_k)^2 \geq 0$$

i.e.,

$$\sum_{k=1}^{n} a_k^2 - 2x \sum_{k=1}^{n} a_k\, b_k + x^2 \sum_{k=1}^{n} b_k^2 \geq 0$$

This is true for all $x \in R$, and $\sum_{k=1}^{n} b_k^2 > 0$, therefore the discriminant of the quadratic in x is non-positive and hence

$$\left(\sum_{k=1}^{n} a_k\, b_k\right)^2 \leq \sum_{k=1}^{n} a_k^2 \sum_{k=1}^{n} b_k^2$$

i.e.,

$$\left|\sum_{k=1}^{n} a_k\, b_k\right| \leq \left(\sum_{k=1}^{n} a_k^2\right)^{1/2} \left(\sum_{k=1}^{n} b_k^2\right)^{1/2}.$$

For the triangle inequality, consider

$$\left[\left(\sum_{i=1}^{n} (x_i - z_i)^2\right)^{1/2} + \left(\sum_{i=1}^{n} (z_i - y_i)^2\right)^{1/2}\right]^2$$

$$= \sum_{i=1}^{n} (x_i - z_i)^2 + \sum_{i=1}^{n} (z_i - y_i)^2 + 2\left(\sum_{i=1}^{n} (x_i - z_i)^2\right)^{1/2} \left(\sum_{i=1}^{n} (z_i - y_i)^2\right)^{1/2}$$

$$\geq \sum_{i=1}^{n} (x_i - z_i)^2 + \sum_{i=1}^{n} (z_i - y_i)^2 + 2\sum_{i=1}^{n} (x_i - z_i)(z_i - y_i)$$

(Using Cauchy-Schwarz inequality)

$$= \sum_{i=1}^{n} [(x_i - z_i) + (z_i - y_i)]^2 = \sum_{i=1}^{n} (x_i - y_i)^2$$

$$\therefore \qquad \left(\sum_{i=1}^{n} (x_i - z_i)^2\right)^{1/2} + \left(\sum_{i=1}^{n} (z_i - y_i)^2\right)^{1/2} \geq \left(\sum_{i=1}^{n} (x_i - y_i)^2\right)^{1/2}$$

i.e., $\quad d(x, y) \leq d(x, z) + d(z, y), \quad \forall\ x, y, z \in \mathbf{R}^n.$

We now generalise $\mathbf{R}^n$ to '*infinite-tuples*' which are sequences, and generalise the above Euclidean metric to the function

$$d(\{x_n\}, \{y_n\}) = \left(\sum_{n=1}^{\infty} (x_n - y_n)^2\right)^{1/2}$$

d is well defined if the series $\sum\limits_{n=1}^{\infty} (x_n - y_n)^2$ converges, and so we must restrict ourselves to those sequences $\{x_n\}$ for which the series $\sum\limits_{n=1}^{\infty} x_n^2$ converges.

Let l_2 be the set of all real sequences $\{x_n\}$, for which the series $\sum\limits_{n=1}^{\infty} x_n^2$ converges, *i.e.*, $\sum\limits_{n=1}^{\infty} x_n^2 < \infty$.

Example 2. Show that the function d defined by

$$d(\{x_n\}, \{y_n\}) = [\sum_{n=1}^{\infty} (x_n - y_n)^2]^{1/2}, \{x_n\}, \{y_n\} \in l_2$$

is a metric on l_2. The metric space (l_2, d) is known as a *Hilbert space*.

■ First we have to show that d is well defined. For this, let $\{x_n\}, \{y_n\} \in l_2$, then by the Cauchy-Schwarz inequality

$$\sum_{k=1}^{n} |x_k||y_k| \leq \sqrt{\left(\sum_{k=1}^{n} |x_k|^2\right)\left(\sum_{k=1}^{n} |y_k|^2\right)}, \quad \forall\ n \in \mathbf{N}$$

$$\therefore \quad \sum_{k=1}^{n} |x_k \; y_k| \le \sqrt{\left(\sum_{k=1}^{n} x_k^2\right)\left(\sum_{k=1}^{n} y_k^2\right)} \le \sqrt{\left(\sum_{k=1}^{\infty} x_k^2\right)\left(\sum_{k=1}^{\infty} y_k^2\right)^2} < \infty$$

This implies $\sum_{n=1}^{\infty} x_n y_n$ converges absolutely and hence converges.

Thus, the series $\sum_{n=1}^{\infty} (x_n - y_n)^2$ being the sum of the three convergent series $\sum_{n=1}^{\infty} x_n^2$, $-2\sum_{n=1}^{\infty} x_n \, y_n$, and $\sum_{n=1}^{\infty} y_n^2$ is also convergent.

Now from the Cauchy-Schwarz inequality we have by taking limits as $n \to \infty$

$$\left| \sum_{k=1}^{\infty} x_k \; y_k \right| \le \sqrt{\left(\sum_{k=1}^{\infty} x_k^2\right)\left(\sum_{k=1}^{\infty} y_k^2\right)} \qquad \text{...(1)}$$

To prove triangle inequality, let $\{a_n\}, \{b_n\}, \{c_n\} \in l_2$.

Put $x_n = a_n - b_n, \; y_n = b_n - c_n, \; \forall \, n \in \mathbf{N}$

Then the triangle inequality in l_2 takes the form:

$$\sqrt{\sum_{n=1}^{\infty} (x_n + y_n)^2} \le \sqrt{\sum_{n=1}^{\infty} x_n^2} + \sqrt{\sum_{n=1}^{\infty} y_n^2}$$

This follows from equation (1), since

$$\begin{aligned}
\sum_{n=1}^{\infty} (x_n + y_n)^2 &= \sum_{n=1}^{\infty} x_n^2 + 2\sum_{n=1}^{\infty} x_n \, y_n + \sum_{n=1}^{\infty} y_n^2 \\
&\le \sum_{n=1}^{\infty} x_n^2 + 2\sqrt{(\Sigma \, x_n^2)\,(\Sigma \, y_n^2)} + \sum_{n=1}^{\infty} y_n^2 \\
&= \left(\left(\sum_{n=1}^{\infty} x_n^2\right)^{\frac{1}{2}} + \left(\sum_{n=1}^{\infty} y_n^2\right)^{\frac{1}{2}}\right)^2.
\end{aligned}$$

Example 3. Show that the set $\mathbf{R}^n$ with d' defined by

$$d'(x, y) = \sum_{i=1}^{n} |x_i - y_i|,$$

is a metric space (d' is called the *rectangular metric* on $\mathbf{R}^n$).

■ For the triangle inequality, consider

$$d'(x, y) = \sum_{i=1}^{n} |x_i - y_i| = \sum_{i=1}^{n} |x_i - z_i + z_i - y_i|$$

$$\leq \sum_{i=1}^{n} |x_i - z_i| + \sum_{i=1}^{n} |z_i - y_i|$$

$$= d'(x, z) + d'(z, y), \quad \forall\ x, y, z \in \mathbf{R}^n$$

Since the metrics d and d' of examples (1) and (3) are different functions, therefore we get different metric spaces $(\mathbf{R}^n, d)$, and $(\mathbf{R}^n, d')$ with the same set $X = \mathbf{R}^n$.

Note that these metrics satisfy the inequality:

$$d(x, y) \leq d'(x, y) \leq \sqrt{n}\ d(x, y).$$

Example 4. Prove that the set $C[a, b]$ of all real-valued functions continuous on the interval $[a, b]$ with the function d defined by

$$d(f, g) = \left(\int_a^b (f(x) - g(x))^2\ dx\right)^{1/2}$$

is a metric space.

To establish triangle inequality we need the following:

- Consider the function, for $t \in [a, b]$

$$\phi(t) = \int_a^b (t\ f(x) + g(x))^2\ dx$$

$$= t^2 \int_a^b f^2(x)\ dx + 2t \int_a^b f(x)\ g(x)\ dx + \int_a^b g^2(x)\ dx$$

Since $\phi(t) \geq 0,\ \forall\ t \in [a, b]$, therefore the discriminant of the quadratic in t should be non-positive, and so

$$\left(\int_a^b f(x)\ g(x)\ dx\right)^2 \leq \int_a^b f^2(x)\ dx \int_a^b g^2(x)\ dx$$

i.e.,

$$\int_a^b f(x)\ g(x)\ dx \leq \left(\int_a^b f^2(x)\ dx\right)^{1/2} \left(\int_a^b g^2(x)\ dx\right)^{1/2} \quad \text{...(1)}$$

Now consider

$$\left[\left(\int_a^b (f(x) - h(x))^2\ dx\right)^{1/2} + \left(\int_a^b (h(x) - g(x))^2\right)^{1/2}\right]^2$$

$$= \int_a^b (f(x) - h(x))^2\ dx + \int_a^b (h(x) - g(x))^2\ dx$$

$$+ 2\left(\int_a^b (f(x) - h(x))^2\ dx\right)^{1/2} \left(\int_a^b (h(x) - g(x))^2\ dx\right)^{1/2}$$

$$\geq \int_a^b (f(x) - h(x))^2\ dx + \int_a^b (h(x) - g(x))^2\ dx$$

$$+ 2\int_a^b (f(x) - h(x))\ (h(x) - g(x))\ dx \quad \text{[using (1)]}$$

$$= \int_a^b (f(x) - h(x) + h(x) - g(x))^2 \, dx = \int_a^b (f(x) - g(x))^2 \, dx$$

Hence

$$\left(\int_a^b (f(x) - g(x))^2 \, dx\right)^{1/2} \le \left(\int_a^b (f(x) - h(x))^2 \, dx\right)^{1/2}$$

$$+ \left(\int_a^b (h(x) - g(x))^2 \, dx\right)^{1/2}.$$

Example 5. Let (X, d) be any metric space. Show that the function d_1 defined by

$$d_1(x, y) = \frac{d(x, y)}{1 + d(x, y)}, \quad \forall \, x, y \in X$$

is a metric on X.

■ For the triangle inequality we proceed as follows:

Using the triangle inequality for the metric d, we have for all $x, y, z \in X$

$$d(x, y) \le d(x, z) + d(z, y)$$

or

$$1 + d(x, y) \le 1 + d(x, z) + d(z, y)$$

or

$$1 - \frac{1}{1 + d(x, y)} \le 1 - \frac{1}{1 + d(x, z) + d(z, y)}$$

or

$$\frac{d(x, y)}{1 + d(x, y)} \le \frac{d(x, z) + d(z, y)}{1 + d(x, z) + d(z, y)}$$

$$\le \frac{d(x, z)}{1 + d(x, z)} + \frac{d(z, y)}{1 + d(z, y)}$$

i.e.,

$$d_1(x, y) \le d_1(x, z) + d_1(z, y).$$

Example 6. *Fréchet Space.* Let X be the set of all sequences of complex numbers. We define the function d by

$$d(x, y) = \sum_{n=1}^{\infty} \frac{1}{2^n} \frac{|x_n - y_n|}{(1 + |x_n - y_n|)}, \quad \forall \, x = \{x_n\}, \; y = \{y_n\} \in X$$

The function d is well defined, since the nth term of the above series is less than $\frac{1}{2^n}$, therefore it is convergent. To prove triangle inequality we first establish the following inequality.

Let $0 \le \alpha \le \beta$, then

$$\alpha + \alpha\beta \le \beta + \alpha\beta$$

Dividing both sides by $(1+\alpha)(1+\beta)$, we obtain

$$\frac{\alpha}{1+\alpha} \le \frac{\beta}{1+\beta} \qquad \ldots(1)$$

Now for any $x = \{x_n\}$, $y = \{y_n\}$ and $z = \{z_n\}$ in X, we have

$$0 \le |x_n - y_n| \le |x_n - z_n| + |z_n - y_n|$$

So from (1) it follows that

$$\frac{|x_n - y_n|}{1+|x_n - y_n|} \le \frac{|x_n - z_n| + |z_n - y_n|}{1+|x_n - z_n| + |z_n - y_n|}$$

$$\le \frac{|x_n - z_n|}{1+|x_n - z_n|} + \frac{|z_n - y_n|}{1+|z_n - y_n|}$$

Multiplying both sides by 2^{-n} and summing w.r.t. n, we get

$$d(x, y) \le d(x, z) + d(z, y)$$

Hence, (X, d) is a metric space.

The very definition of a metric presents the concept of the distance from one point to another. We now define the distance from a point to a set and the distance between two non-empty subsets of a metric space.

For any two non-empty subsets A and B of a metric space (X, d), the *distance* between them denoted by $d(A, B)$ is defined as:

$$d(A, B) = \inf \{d(a, b) \colon a \in A, b \in B\}.$$

If x is a point of X, then the *distance from x to A* denoted by $d(x, A)$ is defined as:

$$d(x, A) = \inf \{d(x, a) \colon a \notin A\}.$$

If $A = \{x \in \mathbf{R} : 0 < x \le 1\}$ and d is the usual metric, then

$$d(0, A) = 0, \text{ although } 0 \notin A.$$

Similarly $d(A, B) = 0$ does not imply that A and B have common elements, as can be seen by the following example.

Example 7. Let $X = \left\{1, \frac{1}{2}, \frac{1}{3}, \ldots, \frac{1}{n}, \ldots\right\}$ and d is the usual metric defined on X.

Let $A = \left\{1, \frac{1}{3}, \frac{1}{5}, \ldots, \frac{1}{2n-1}, \ldots\right\}$, and

$$B = \left\{\frac{1}{2}, \frac{1}{4}, \frac{1}{6}, \ldots, \frac{1}{2n}, \ldots\right\}$$

Then $d(A, B) = 0$, although $A \cap B = \phi$.

1.1 Diameter of a Non-Empty Set

Definition. The *diameter* of any non-empty subset $A \subseteq X$ denoted by $d(A)$ is defined as:

$$d(A) = \sup \{d(a, b) : a, b \in A\}.$$

If $d(A) < \infty$, then the diameter of A is said to be finite otherwise infinite.

By convention $d(\phi) = -\infty$.

Definition. A metric d on a non-empty set X is said to be *bounded* if there exists a real number $k > 0$ such that

$$d(x, y) \leq k, \quad \forall\, x, y \in X$$

i.e.,

$$d(X) \leq k,$$

(X, d) is then called a *bounded metric space*, otherwise unbounded.

Example 8. Let $\mathbf{R}_\infty$ be the *extended set* of real numbers (*i.e.*, the set of real numbers including $-\infty$ and $+\infty$).

The function d defined by

$$d(x, y) = |\, f(x) - f(y)\,|, \quad \forall\, x, y \in \mathbf{R}_\infty$$

where $f(x)$ is given by

$$f(x) = \begin{cases} \dfrac{x}{1 + |x|}, & \text{when } -\infty < x < \infty \\ 1, & \text{when } x = \infty \\ -1, & \text{when } x = -\infty \end{cases}$$

Show that $(\mathbf{R}_\infty, d)$ is a bounded metric space.

■ For the triangle inequality

$$d(x, y) = \left| \frac{x}{1 + |x|} - \frac{y}{1 + |y|} \right|$$

$$= \left| \frac{x}{1 + |x|} - \frac{z}{1 + |z|} + \frac{z}{1 + |z|} - \frac{y}{1 + |y|} \right|$$

$$\leq \left| \frac{x}{1 + |x|} - \frac{z}{1 + |z|} \right| + \left| \frac{z}{1 + |z|} - \frac{y}{1 + |y|} \right|$$

$$= d(x, z) + d(z, y), \quad \forall\, x, y, z \in \mathbf{R}$$

If $x = \infty$, $y = -\infty$, then

$$d(x, y) = |1 - (-1)| \le \left|1 - \frac{z}{1 + |z|}\right| + \left|\frac{z}{1 + |z|} - (-1)\right|$$

$$= d(x, z) + d(z, y)$$

Similarly when $x = -\infty$, $y = +\infty$, the triangle inequality holds.

Hence $(\mathbf{R}_\infty, d)$ is a metric space.

Moverover, if x and y are two elements of $\mathbf{R}_\infty$, then

$$-1 \le f(x) \le 1, \text{ and } -1 \le f(y) \le 1$$

$\therefore$ $$d(x, y) = |f(x) - f(y)| \le 2, \quad \forall\, x, y \in \mathbf{R}_\infty$$

Hence $(R_\infty\, d)$ is a bounded metric space.

Remark: It is to be noted that even when a metric space is unbounded, we can define another metric in many ways, so that the resulting metric space is bounded. As from example 5 if (X, d) is any metric space, then (X, d_1) is a bounded metric space, where

$$d_1(x, y) = \frac{d(x, y)}{1 + d(x, y)}, \; x, y \in X$$

Since $$0 \le d_1(x, y) < 1, \quad \forall\, x, y \in X.$$

EXERCISE

1. Let X be a non-empty set and a function d from $X \times X$ into $\mathbf{R}$ satisfies :
   (*i*) $d(x, y) = 0$, if and only if $x = y$, and
   (*ii*) $d(x, y) \le d(x, z) + d(y, z)$, $\forall\, x, y, z \in X$.

   Prove that (X, d) is a metric space.

   [*Hint*: Take $y = x$, in (*ii*), $d(x, z) \ge 0$. Take $x = z$ in (*ii*), $d(z, y) \le d(y, z)$ and interchange the role of y and z.]
2. Show that the conditions
   (*i*) $d(x, y) = 0$, if and only if $x = y$, and
   (*ii*) $d(x, y) \le d(x, z) + d(z, y)$, $\forall\, x, y, z \in X$

   are not sufficient to ensure that the function $d : X \times X \to \mathbf{R}$ is a metric on a non-empty set X.
3. Give an example of a function $d : X \times X \to R$ defined on a non-empty set X satisfying the following three conditions but not a metric on X.
   (*i*) $d(x, y) \ge 0$, and $x = y \Rightarrow d(x, y) = 0$
   (*ii*) $d(x, y) = d(y, x)$, $\forall\, x, y \in X$
   (*iii*) $d(x, y) \le d(x, z) + d(z, y)$, $\forall\, x, y, z \in X$

   [*Hint*: Take $X = \mathbf{R}^2$, $d(x, y) = |x_1 - y_1|$, where $x = (x_1, x_2)$, $y = (y_1, y_2)$. Such a metric is called a *Pseudometric*.]

4. Prove that if (X, d) is a metric space, then

$$|d(x, z) - d(y, z)| \le d(x, y), \quad \forall\ x, y, z \in X.$$

[*Hint*: Apply triangle inequality to $d(x, z)$ and $d(y, z)$ separately.]

5. Prove that, if (X, d) is a metric space, and $x_1, x_2, x_3, \dots x_n \in X$, then

$$d(x_1, x_n) \le d(x_1, x_2) + d(x_2, x_3) + d(x_3, x_4) + \dots + d(x_{n-1}, x_n).$$

6. Functions d_1 and d_2 from $\mathbf{R} \times \mathbf{R} \to \mathbf{R}$ are defined by

$$d_1(x, y) = \exp\{|x - y|\}, \text{ and } d_2(x, y) = \max\{x - y, 0\}.$$

Is either d_1 or d_2 a metric on $\mathbf{R}$?

7. Prove that the function $d^*: \mathbf{R}^n \times \mathbf{R}^n \to \mathbf{R}$ defined by

$$d^*(x, y) = \max_{1 \le i \le n} |x_i - y_i|, \quad \forall\ x = (x_1, x_2 \dots x_n),\ y = (y_1, y_2 \dots y_n) \in \mathbf{R}^n$$

is a metric on $\mathbf{R}^n$. Also prove that

$$d^*(x, y) \le d\,(x, y) \le \sqrt{n} d^*(x, v), \quad \forall\ x, y \in \mathbf{R}^n$$

where d is the *Euclidean metric* on $\mathbf{R}^n$.

8. Consider the set l_p of all sequences $\{x_n\}$ of complex numbers satisfying the convergence condition $\sum_{n=1}^{\infty} |x_n|^p < \infty$, for some fixed $p \ge 1$, where the distance between points is defined by

$$d(x, y) = \left(\sum_{n=1}^{\infty} |x_n - y_n|^p\right)^{1/p}, \quad \forall\ x = \{x_n\},\ y = \{y_n\} \in l_p$$

Show that (l_p, d) is a metric space.

[*Hint*: $d(x, y)$ is well defined, can be seen by using, *Minkowski's inequality*:

$$\left(\sum_{n=1}^{\infty} |x_n + y_n|^p\right)^{1/p} \le \left(\sum_{n=1}^{\infty} |x_n|^p\right)^{1/p} + \left(\sum_{n=1}^{\infty} |y_n|^p\right)^{1/p}, \quad \forall\ \{x_n\}, \{y_n\} \in l_p]$$

9. Let H_∞ denote the set of all real sequences $\{x_n\}$ such that $|x_n| \le 1,\ \forall\ n \in N$, then prove that the function d defined by

$$d(\{x_n\}, \{y_n\}) = \sum_{n=1}^{\infty} \frac{|x_n - y_n|}{2^n}, \{x_n\}, \{y_n\} \in H_\infty$$

is a metric on H_∞ [(H_∞, d) is called the *Hilbert-cube*].

10. Let $(X, d_1), (X, d_2)$ be two metric spaces and k is a positive real number. Which of the following are metric spaces:

$$(X, kd_1), (X, \sqrt{d_1}), (X, d_1^2), (X, \min(1, d_1)), (X, d_1 d_2),$$
$$(X, d_1 + d_2), (X, \max(d_1, d_2)), (X, \min(d_1, d_2))?$$

11. Prove that the function $d: \mathbf{C} \times \mathbf{C} \to \mathbf{R}$ defined by

$$d(x, y) = \frac{2|x - y|}{\sqrt{1 + |x^2|}\sqrt{1 + |y|^2}}$$

is a metric on the set of all complex numbers.

12. Let $R[0, 1]$ be the set of all Riemann integrable functions defined on $[0, 1]$ and let d be defined by

$$d(f, g) = \int_0^1 |f(x) - g(x)|\,dx, \quad \forall\, f, g \in R[0, 1],$$

Show that d is not a metric on $R[0, 1]$.

13. If d is a metric on X, then show that min $\{d(x, y), 1\}$ is a bounded metric on X.

14. If $d : \mathbf{N} \times \mathbf{N} \to \mathbf{R}$ is defined by, for $m, n \in \mathbf{N}$,

$$d(m, n) = 0, \text{ if } m = n, \text{ and for } m \neq n,\ d(m, n) = 1/5^k,$$

where $m - n = 5^k r$, are r is not a multiple of 3. Prove that (N, d) is a bounded metric space.

15. Let $C'[a, b]$ denote the set of all real-valued functions defined on the closed interval $[a, b]$ having continuous first order derivatives on $[a, b]$. Prove that the function d defined on $C'[a, b]$ by

$$d(f, g) = \sup_{a \le x \le b} |f(x) - g(x)| + \sup_{a \le x \le b} |f'(x) - g'(x)|, \ \forall\, f, g \in C'[a, b]$$

is a metric on $C'[a, b]$.

16. Let A and B be any two non-empty subsets of a metric space X.
Prove that
   (*i*) If $A \subseteq B$, then $d(A) \le d(B)$,
   (*ii*) $d(A \cup B) \le d(A) + d(B) + d(A, B)$
   (*iii*) If $A \cap B \neq \phi$, then $d(A \cup B) \le d(A) + d(B)$.

2. OPEN AND CLOSED SETS

In this section we shall study the concept of neighbourhoods, open sets and closed sets in a metric space (X, d) and develop some of the important results relating to these concepts. We begin by defining open spheres and closed spheres.

2.1 Open and Closed Spheres

Let (X, d) be any metric space, and $a \in X$. Then for any $r > 0$, the set

$$S_r(a) = \{x \in X : d(x, a) < r\}$$

is called an *open sphere* (or *open ball*) of radius r centered at a.

The set

$$S_r[a] = \{x \in X : d(x, a) \le r\}$$

is called a *closed sphere* of radius r centered at a. It is clear that

$$S_r(a) \subset S_r[a]$$

for every $a \in X$, and for every $r > 0$.

ILLUSTRATIONS

1. In the metric space $(\mathbf{R}, d)$ of real numbers with the usual metric d, the open sphere $S_r(a)$ is the open interval $]a - r, a + r[$, and the closed sphere $S_r[a]$ is the closed interval $[a - r, a + r]$, where $a \in \mathbf{R}$, and $r > 0$.

2. In the discrete space (X, d), the open sphere $S_r(a)$ for $a \in X$ is given by

$$S_r(a) = \begin{cases} \{a\}, & \text{if } 0 < r \le 1 \\ X, & \text{if } r > 1 \end{cases}$$

and the closed sphere $S_r[a]$ is given by

$$S_r[a] = \begin{cases} \{a\}, & \text{if } 0 < r < 1 \\ X, & \text{if } r \ge 1. \end{cases}$$

3. The open sphere in the complex plane is the inside of a circle with center at a and radius r.
4. The open sphere $S_r(f_0)$ in the metric space $C[a, b]$ of all real-valued continuous functions defined on $[a, b]$ is a strip of width $2r$ centered on the graph of f_0.

Example 9. Let (X, d) be a metric space and $S_r(x)$ the open sphere with centre x and radius r. Let A be a subset of X with diameter less than r, which intersects $S_r(x)$, then $A \subseteq S_{2r}(x)$.

■ Since $A \cap S_r(x) \neq \phi$, let $a \in A \cap S_r(x)$

then $$d(a, x) < r, \quad a \in A, \quad r > 0.$$

Let y be an arbitrary element of A then by triangle inequality

$$\begin{aligned} d(y, x) &\le d(y, a) + d(a, x) \qquad [\because\ y, a \in A \text{ and } d(A) < r] \\ &< r + r = 2r \end{aligned}$$

This implies $y \in S_{2r}(x)$.

2.2 Neighbourhood of a Point

Let (X, d) be a metric space and $a \in X$. A subset N_a of X is called a *neighbourhood* of a point $a \in X$, if there exists an open sphere $S_r(a)$ centered at a and contained in N_a; *i.e.*, $S_r(a) \subseteq N_a$, for some $r > 0$.

Example 10. Every open sphere is a neighbourhood of each of its points.

■ Let $S_r(a)$ be an open sphere, and $x \in S_r(a)$. If $x = a$, then $a \in N_a \subset S_r(a)$. Therefore suppose that $x \neq a$. In order to show that $S_r(a)$ is a neighbourhood of x, we must show that there exists $r_1 > 0$ such that

$$S_{r_1}(x) \subseteq S_r(a)$$

Now $x \in S_r(a)$ implies $d(x, a) < r$. Take $r_1 = r - d(x, a)$.

Then $y \in S_{r_1}(x)$ implies by using triangle inequality

$$d(y, a) \le d(y, x) + d(x, a) < r_1 + d(x, a) = r$$

i.e., $$y \in S_r(a)$$

Hence $$S_{r_1}(x) \subseteq S_r(a).$$

2.3 Open Set

Definition. A subset G of a metric space (X, d) is said to be **open** in X with respect to the metric d, if G is a neighbourhood of each of its points. In other words, if for each $a \in G$, there is an $r > 0$ such that

$$S_r(a) \subseteq G.$$

ILLUSTRATIONS

1. The empty set ϕ and the entire space X with any metric are open sets.
2. Every open sphere is an open set.
3. Let A be the annulus consisting of the complex numbers z such that $1 < |z| < 2$ with the usual metric d, then A is open.
4. The subset $S = \{(x, y): x^2 + y^2 < 1,\ x, y \in \mathbf{R}\}$ of $\mathbf{R}^2$ with the Euclidean metric is an open set.
5. The subset $A = \{(x, y): y^2 < x;\ x, y \in \mathbf{R}\}$ of $\mathbf{R}^2$ with the Euclidean metric is an open set.
6. Let $A = \{f \in C[a, b]: \inf\limits_{x \in [a, b]} f(x) > 0\}$. Then A is open with respect to the sup metric in $C[a, b]$.

Example 11. Show that every set in a discrete space (X, d) is open.

■ Let G be any non-empty subset of the discrete space (X, d) and x be any point of G. Then the open sphere $S_r(x)$ with $r \leq 1$ is the singleton set $\{x\}$ which is contained in G *i.e.*, each point of G is the centre of some open sphere contained in G. In particular, each singleton set is open.

Example 12. Show that on the real line with the usual metric the singleton set $\{x\}$ is not open.

■ For the metric space $(\mathbf{R}, d)$ each open sphere $S_r(x)$ is the bounded open interval $]x - r, x + r[$ and for no value of r (how so-ever small it may be) this sphere is contained in $\{x\}$.

Hence $\{x\}$ is not open in $(\mathbf{R}, d)$.

Remark: It is to be observed that a given subset of a metric space may be open with respect to one metric but may not be open with respect to another metric as can be seen by the following example.

Example 13. Let $\mathbf{R}$ be the set of reals, d the usual metric and d' the discrete metric on the same set $\mathbf{R}$. Then show that every singleton set $\{x\},\ x \in \mathbf{R}$ is open in $(\mathbf{R}, d')$ but not so in $(\mathbf{R}, d)$.

■ Every singleton set $\{x\},\ x \in \mathbf{R}$ is open in $(\mathbf{R}, d')$ being an open sphere $S_r(x),\ r < 1$, *i.e.*, bounded open interval $]x - r, x + r[$ for $r < 1$ contains only one point of the set, x of the space. But $\{x\}$ is not open in $(\mathbf{R}, d)$. Because every open sphere $S_r(x),\ r < 1,$ is the bounded open interval $]x - r, x + r[$ $\not\subset \{x\}$.

Example 14. Show that the subset $A = [0, 1[$ of the metric space (X, d), where $X = [0, 2[$, and d is the usual metric, is an open set.

■ Let $x \in A = [0, 1[$.

If $x = 0$, then $S_{1/2}(0) = [0, \frac{1}{2}[\subseteq A$. If $x \neq 0$, choose $r = \min\{x, 1 - x\}$. Clearly $r > 0$ and

$$S_r(x) =]x - r, x + r[\subseteq [0, 1[= A.$$

Hence, A is open in X.

Definition. Two metrics d and d' on the same set X are said to be *equivalent*, if every set open in (X, d) is open in (X, d'), and vice versa.

Example 15. Let (X, d) be any metric space and let

$$d(x, y) = \frac{d(x, y)}{1 + d(x, y)}, \quad \forall\ x, y \in X$$

Show that d and d' are equivalent.

■ We have already shown that d' is a metric on X (Example 5).

Let G be any open subset of (X, d). Then for each $x \in G, \exists$ an open sphere,

$$S_r(x) = \{y \in X : d(y, x) < r\} \subseteq G$$

Let $r_1 = \dfrac{r}{1+r}$, then $r_1 < r$.

Now,

$$\{y \in X : d'(y, x) < r_1\} \subseteq \{y \in X : d(y, x) < r\}$$

$$d'(y, x) < r_1 \Rightarrow \frac{d(y, x)}{1 + d(y, x)} < \frac{r}{1+r}.$$

Thus every point of G is the centre of some open sphere in (X, d') contained in G.

Consequently, every set open in (X, d) is open in (X, d'). Again, let G be any open set in (X, d') so, $\exists$ an open sphere

$$S_r(x) = \{y \in X : d'(y, x) < r\} \subseteq G$$

Since $d'(x, y) < 1$, we may assume $r < 1$.

Let $r' = \dfrac{r}{1-r}$.

Now each point of G is the centre of an open sphere contained in G implying that G is open in (X, d).

Ex. Let (X, d) be a metric space, and let $d'(x, y) = \min\{1, d(x, y)\}$ for all $x, y \in X$. Then show that d and d' are equivalent.

Theorem 1. *In any metric space (X, d),*

(*i*) *the union of an arbitrary family of open sets is open,*

(*ii*) *the intersection of a finite number of open sets is open.*

(*i*) Let $\{G_\alpha : \alpha \in \Lambda\}$ be an arbitrary family of open sets in X, where Λ is any non-empty index set.

Let $G = \underset{\alpha \in \Lambda}{\cup} G_\alpha$.

If $G = \phi$ then G is open. Suppose $G \neq \phi$. Let x be any element of G. Since $G = \underset{\alpha \in \Lambda}{\cup} G_\alpha$, therefore there is an $\alpha_0 \in \Lambda$ such that $x \in G_{\alpha_0}$.

The set G_{α_0} being open implies that there exists $r > 0$, such that

$$S_r(x) \subseteq G_{\alpha_0},$$

and so

$$S_r(x) \subseteq \underset{\alpha \in \Lambda}{\cup} G_\alpha \qquad (\because\ G_{\alpha_0} \subseteq \underset{\alpha \in \Lambda}{\cup} G_\alpha)$$

Hence, G is an open set.

(*ii*) Let $G_1, G_2, ..., G_n$ be any finite number of open sets in X, and

$$G = \bigcap_{i=1}^{n} G_i.$$

If $G = \phi$, then G is open. Suppose $G \neq \phi$.

Let $x \in G = \bigcap_{i=1}^{n} G_i$, then $x \in G_i$, $\forall\, i = 1, 2, ...n$

Since each $G_i, i = 1, 2 ... n$, is open $\exists\, r_i > 0$, such that

$$S_{r_i}(x) \subseteq G_i, \quad \forall\, i = 1, 2, ...n$$

Let $r = \min\{r_1, r_2, ..., r_n\}$.

Then $r > 0$ and

$$S_r(x) \subseteq S_{r_i}(x) \subseteq G_i \quad \forall\, i = 1, 2, ...n$$

This implies

$$S_r(x) \subseteq \bigcap_{i=1}^{n} G_i = G$$

Hence, G is open in X.

Remarks:

1. The part (*ii*) of the above theorem need not be true for the intersection of an arbitrary family of open sets.

   For example, consider a family of open sets $G_n =]-1/n, 1/n[, n \in \mathbf{N}$ in $(\mathbf{R}, d)$, where d is the usual metric. The intersection

   $\bigcap_{n=1}^{\infty} G_n = \{0\}$, which is not open in **R**.

   Definition. If X is any set, and **F** is a collection of subsets of X satisfying,

   (*i*) $\phi, X \in \mathbf{F}$

   (*ii*) The union of an arbitrary family of sets in **F** is a member of **F**.

   (*iii*) The intersection of a finite number of sets in **F** is a number of **F**, then **F** is called a *topology* for X.

   For example, the collection of all open subsets of a metric space X is a topology for X.

2. We note that, just as every open set on the real line is a union of open intervals, similarly every open set in the metric space (X, d) can be written as a union of open spheres. If $G \subseteq X$ is open, then for each $x \in G, \exists\, r_x > 0$ such that

$$S_{r_x}(x) \subseteq G.$$

Hence $G = \bigcup_{x \in G} S_{r_x}(x)$.

Conversely, the union of open spheres is always an open set in (X, d) (the first part of the above theorem). However, in the case of real line an explicit description of open sets can be given.

Lemma. *Let* **F** *be the family of open intervals in* **R**, *no two of which are disjoint, then* $\bigcup_{\alpha \in \Lambda} I$ *is an open interval.*

Let $I_0 = \bigcup_{I \in F} I$, and let $a, b \in I_0$, and $a < c < b$.

CHAPTER 19

Then $a \in I_1$ and $b \in I_2$, for some $I_1, I_2 \in \mathbf{F}$.

Let $I_1 =]a_1, b_1[$ and $I_2 =]a_2, b_2[$, then

$$a_1 < a < c < b < b_2 \qquad ...(1)$$

If $b_1 \leq a_2$, then $I_1 \cap I_2 = \phi$, which is impossible, so $b_1 > a_2$. Then either $c < b_1$ or $c \geq b_1$. In the former case $c \in I_1$, and in the latter $c \in I_2$ ($\because a_2 < b_1$) and consequently in either case $c \in I_0$. Thus I_0 is an interval. The interval I_0 must be an open interval because it is an open set (being union of open intervals).

Theorem 2. *Every non-empty open set on the real line is the union of a countable collection of pairwise disjoint open intervals.*

Let G be a non-empty open subset of **R**. For each $x \in G$, let I_x be the union of all the open intervals which contain x, and are contained in G (such intervals exist because G is open). By the above Lemma, each I_x is an open interval. Obviously I_x contains every open interval which contains x and is contained in G. Moreover

$$G = \bigcup_{x \in G} I_x$$

We shall show that any two members in the above union are either disjoint or identical. For this let $x, y \in G$, and suppose that

$$I_x \cap I_y \neq \phi$$

Then, by the above lemma, the set $I_x \cup I_y$ is an open interval which contains both x and y. Therefore by definition of I_x and I_y it follows that

$$I_x \cup I_y \subseteq I_x, \text{ and } I_x \cup I_y \subseteq I_y.$$

Consequently $I_x = I_y$.

Let **F** be the collection of all distinct sets of the form I_x with $x \in G$. This being a disjoint collection of open intervals whose union is in G.

Now it remains to show that **F** is countable. The set $Q \cap G$ of all rational numbers in G is countable.

We define a function f from $Q \cap G$ into **F** as follows:

Given $r \in Q \cap G$, let $f(r)$ be the unique set I_r in **F** that contains r (the set I_r is unique because the sets in **F** are pairwise disjoint).

f is obviously onto, since each open interval in **R** contains a rational number. Hence **F** is countable.

2.4 Limit Points

Definition. Let A be any subset of a metric space (X, d). A point 'a' of X is called an *adherent point* of A, if every open sphere centered at 'a' contains a point of A.

Adherent points are of two types:

(*i*) isolated points,

(*ii*) limit points.

An adherent point 'a' of a subset A of X is called an *isolated point* if every open sphere centered at 'a' contains no point of A other than a itself.

An adherent point 'a' of a subset A of X is said to be a *limit point* of A if every open sphere centered at 'a' contains at least one member of A other than a

i.e., $$S_r(a) \cap (A - \{a\}) \neq \phi, \quad \forall\, r > 0$$

The essential idea here is that points of A different from 'a' get arbitrarily close to a or 'pile up' at a.

The limit point is also known as a *cluster point*, a *condensation point* or an *accumulation point*. If 'a' is a limit point of A then every open sphere centered at a contains infinitely many elements of A, and conversely. The limit point of A may or may not be a member of A.

Derived Set. The set of all limit points of A is called the *derived set* of A and is denoted by A'.

ILLUSTRATIONS

1. Let **R** be the set of reals with the usual metric d. Let $A = [0, 1[$. Every point of A is a limit point of A. Further '1' is also a limit point of A which is not a member of A. Here $A' = [0, 1]$.
2. Let $A = \{1, \frac{1}{2}, \frac{1}{3} \ldots\}$ with the usual metric in **R**. '0' is the only limit point of A which is not a member of A, so that $A' = \{0\}$.
3. The derived set of every subset of a discrete space is empty.
4. Every real number is a limit point of the set of rationals.
5. The set of integers has no limit point.
6. A finite set has no limit point.
7. Let A be the annulus consisting of complex numbers z such that $1 < |z| < 2$, with the usual metric, then $A' = \{z : 1 \leq |z| \leq 2\}$.

2.5 Closed Sets

Definition. A subset F of a metric space (X, d) is said to be closed if F contains all its limit points.

ILLUSTRATIONS

1. The empty set ϕ and the whole space X are closed sets in every metric space (X, d).
2. On the real line with the usual metric the set **N** of natural numbers is closed.
3. The set **Q** of rational numbers is not closed.
4. Every closed interval on the real line is a closed set.
5. Every finite subset of a metric space is closed.
6. Let $A = \{f \in C[a, b] : f(a) = 1\}$. Then A is closed with respect to the sup metric in $C\,[a, b]$.
7. The set $\left\{\dfrac{1+i}{n} : n \in N\right\}$ is neither open nor closed with respect to the usual metric in the complex plane.

Theorem 3. *Let (X, d) be any metric space. A subset F of X is closed if and only if its complement in X is open.*

Suppose F is closed. If $F^c = X - F = \phi$, then it is open.

Assume that $X - F \neq \phi$. Let $x \in X - F$; then $x \notin F$. F being closed implies x is not a limit point of F and so there exists $r > 0$, such that

$$S_r(x) \cap F = \phi$$

i.e,
$$x \in S_r(x) \subseteq X - F$$

Hence F^c is open.

Conversely, suppose that $X - F$ is open. If $x \in X - F$, then there exists $r > 0$, such that

$$S_r(x) \subseteq X - F$$

i.e., $S_r(x) \cap F = \phi \Rightarrow x$ cannot be a limit point of F.

Since $x \in X - F$ is arbitrary, therefore F does not have any limit point outside it. Consequently F is closed.

From the above theorem we can easily see that every subset of a discrete space is closed. Since in a discrete space (X, d) every subset of it is open so if F is an arbitrary subset of X, then $X - F$ being a subset of X is open and therefore F is closed.

Example 16. Every closed sphere is a closed set.

■ Let $S_r[x]$ be any closed sphere in a metric space (X, d). If $X - S_r[x] = \phi$, then ϕ is open.

Assume $X - S_r[x] \neq \phi$. Let $y \in X - S_r[x]$. Then $y \notin S_r[x]$.

This implies $d(y, x) > r$. Let $r_1 = d(y, x) - r$.

The open sphere $S_{r_1}(y) \subseteq X - S_r[x]$, for if $z \in S_{r_1}(y)$, then $d(z, y) < r_1$, and so

$$d(z, y) < d(y, x) - r$$

i.e.,
$$r < d(y, x) - d(z, y) = d(x, z) \quad \text{(by triangle inequality)}$$

Thus $z \in S_{r_1}(y) \subseteq X - S_r[x]$

This implies $X - S_r[x]$ is open. Hence $S_r[x]$ is closed.

Theorem 4. *In any metric space* (X, d),

(*i*) *the intersection of an arbitrary family of closed sets is closed*,

(*ii*) *the union of a finite number of closed sets is closed*.

(Proofs follow from Theorems 6 and 7 of Chapter 2.)

[For alternative proof. *Hint* : (*i*) Let $\{F_\alpha : \alpha \in \Lambda\}$ be any family of closed sets, then $\bigcap_{\alpha \in \Lambda} F_\alpha$ is closed, $\because X - \bigcap_{\alpha \in \Lambda} F_\alpha = \bigcup_{\alpha \in \Lambda} (X - F_\alpha)$, Λ is any index set.]

Remark: The arbitrary union of closed sets in a metric space is not necessarily a closed set.
For example, consider the family $\{F_n : n \in \mathbf{N}\}$, where $F_n = [1/n, 1]$ is closed for each $n = 1, 2, \ldots$,

$$\bigcup_{n=1}^{\infty} F_n =]0, 1],$$

which is not closed.

2.6 Subspaces

Definition. Let (X, d) be a metric space. Let Y be a non-empty subset of X. Then the restriction map d_Y of the metric d to $Y \times Y$ is a metric for Y called the *induced metric* and the metric space (Y, d_Y) is called a *subspace* of (X, d).

The closed unit interval $[0, 1]$ and the set of all rational numbers are subspaces of **R**; and the unit circle, the closed unit disc and the open unit disc are subspaces of the space $(\mathbf{C}, d)$ of complex numbers. In fact the real line itself is a subspace of the space complex numbers.

If $Y \subseteq X$, (X, d) is a metric space, and $y \in Y$, then we shall denote the open sphere centered at y with radius r in (Y, d_Y) by $S_r^Y(y)$.

i.e.,
$$S_r^Y(y) = \{x \in Y : d_Y(x, y) < r\}$$

It is easy to verify that

$$S_r^Y(y) = S_r(y) \cap Y.$$

From this it follows that a subset of Y which is open in X is also open in Y. However, the converse may not be true as can be seen by the following examples.

1. Take $Y = [0, 1]$, $X = \mathbf{R}$, d the usual metric, then $S_{\frac{1}{2}}^Y(0) = [0, \frac{1}{2}[$ is open in Y, but not in **R**. Note that Y is not open in **R**.
2. An open interval of the real line is not an open subset of the complex plane.

The following theorem gives a criterion for a subset to be open in a subspace.

Theorem 5. *Let (X, d) be a metric space and $Y \subseteq X$, then a subset A of Y is open in (Y, d_Y) if and only if there exists a set G open in (X, d) such that*

$$A = G \cap Y.$$

Assume that A is open in (Y, d_Y). Then for each $a \in A$ there exists $r_a > 0$ such that

$$S_{r_a}^Y(a) \subseteq A$$

so that

$$A = \bigcup_{a \in A} S_{r_a}^Y(a)$$

But since

$$S_{r_a}^Y(a) = S_{r_a}(a) \cap Y$$

$$\therefore \quad A = \bigcup_{a \in A} (S_{r_a}(a) \cap Y) = G \cap Y,$$

where $$G = \bigcup_{a \in A} S_{r_a}(a) \text{ is open in } (X, d).$$

Conversely, suppose that there is a set G which is open in (X, d) with $G \cap Y = A$.

Let $a \in A$, then $a \in G$, and so there exists $r > 0$, such that

$$S_r(a) \subseteq G$$

This implies

$$S_r^Y(a) = S_r(a) \cap Y \subseteq G \cap Y = A$$

i.e., $$S_r^Y(a) \subseteq A$$

Hence, A is open in (Y, d_Y).

From the above theorem, it follows that an open subset of Y is open in X if and only if Y itself is open in X.

A similar criterion for closed sets is the following:

Theorem 6. *Let (X, d) be a metric space and $Y \subseteq X$, then a subset A of Y is closed in (Y, d_Y) if and only if there exists a set F closed in (X, d) such that*

$$A = F \cap Y.$$

(Proof follows from the above theorem by taking complements.)

2.7 Closure of a Set

Definition. Let A be any subset of a metric space (X, d). The *closure of A* denoted by $\overline{A}$ is the set of all adherent points of A.

i.e., $$\overline{A} = A \cup A'$$

Symbolically $$\overline{A} = \{x \in X : S_r(x) \cap A \neq \phi, \text{ for all } r > 0\}.$$

Properties:

Let A and B be any two subsets of a metric space (X, d). Then

(1) $\overline{A}$ is a closed set.

(2) If $A \subseteq B$, then $\overline{A} \subseteq \overline{B}$.

(3) $\overline{A}$ is the smallest closed superset of A.

(4) $A = \overline{A}$ if and only if A is closed.

(5) $\overline{A}$ is the intersection of all closed sets containing A.

(6) $\overline{A \cup B} = \overline{A} \cup \overline{B}$

(7) $\overline{A \cap B} \subseteq \overline{A} \cap \overline{B}$

(1) In order to show that $\overline{A}$ is a closed set we shall show that its complement $(\overline{A})^c$ is open.

If $(\overline{A})^c = \phi$ then ϕ is open. Assume that $(\overline{A})^c \neq \phi$.

Let $x \in (\overline{A})^c$, then $x \notin \overline{A} \Rightarrow$ there exists at least one $r > 0$ such that

$$S_r(x) \cap A = \phi.$$

Now to show $S_r(x) \cap \overline{A} = \phi$, we take a $y \in S_r(x)$, then $d(y, x) < r$.

Let $r_1 = r - d(y, x)$.

Clearly $r_1 > 0$ and $S_{r_1}(y) \subseteq S_r(x)$

$\Rightarrow \quad S_{r_1}(y) \cap A = \phi$, for at least one r_1 $[\because S_{r_1}(y) \cap A \subseteq S_r(x) \cap A]$

$\Rightarrow \quad y \notin \overline{A}$

Since y is an arbitrary member of $S_r(x)$, therefore

$S_r(x) \subseteq (\overline{A})^c$. This implies $(\overline{A})^c$ is open.

Hence $\overline{A}$ is open.

(2) Let $x \in \overline{A}$ then

$$S_r(x) \cap A \neq \phi, \text{ for all } r > 0$$

This implies

$$S_r(x) \cap B \neq \phi, \quad (\because A \subseteq B)$$

i.e., $\quad x \in \overline{B}$.

Hence, $\overline{A} \subseteq \overline{B}$.

(3) We know that $\overline{A}$ is a closed set, and $A \subseteq \overline{A}$. To show that $\overline{A}$ is the smallest closed set containing A, we assume that if F is any other closed set containing A, then $A \subseteq F \Rightarrow \overline{A} \subseteq \overline{F} = F$ [$\because F$ is closed]. Since F is arbitrary, so $\overline{A}$ is the smallest closed set containing A.

(4) If $A = \overline{A}$, then by (1) $\overline{A}$ is closed, and so A is closed.

Conversely, let A be any closed set.

Since $A \subseteq \overline{A}$, so we need to show that $\overline{A} \subseteq A$.

Let x be any element of $\overline{A}$, then either $x \in A$ or $x \notin A$.

If $x \in A$, then the result is proved.

If $x \notin A$, and $x \in \overline{A}$, then for every $r > 0$, the open sphere $S_r(x)$ contains a point of A other than x.

$\Rightarrow \quad x$ is a limit point of A.

But A being closed, therefore x must belong to A. Hence $\overline{A} \subseteq A$.

(5) Let F be the intersection of all closed sets containing A. Then F is closed.

$$A \subseteq F \Rightarrow \overline{A} \subseteq \overline{F} = F$$

i.e., $\quad \overline{A} \subseteq F$,

Thus every closed set which contains A, contains $\overline{A}$. But $\overline{A}$ is a closed set containing A. F, being the intersection of all closed sets containing A, is contained in $\overline{A}$. Therefore $\overline{A} = F$.

(6) We know that

$$A \subseteq A \cup B, \text{ and } B \subseteq A \cup B$$

$\therefore \quad \overline{A} \subseteq \overline{A \cup B}$, and $\overline{B} \subseteq \overline{A \cup B}$

And so $\quad \overline{A} \cup \overline{B} \subseteq \overline{A \cup B}$.

Now to show that

$$\overline{A \cup B} \subseteq \overline{A} \cup \overline{B}$$

We proceed as follows:

Let, if possible $x \in \overline{A \cup B}$, but $x \notin \overline{A} \cup \overline{B}$. The x is neither an adherent point of A nor that of B. Consequently, there exist open spheres $S_{r_1}(x)$, and $S_{r_2}(x)$ containing no point of A and B respectively.

CHAPTER 19

Take $r = \min\{r_1, r_2\}$, then

$S_r(x)$ contains no point of A as well as no point of B, and therefore of $A \cup B$.

$\therefore$ x is not an adherent point of $A \cup B$.

i.e., $x \notin \overline{A \cup B}$.

Thus we arrive at a contradiction.

Hence $x \in \overline{A \cup B} \Rightarrow x \in \overline{A} \cup \overline{B}$.

(7) Since $A \cap B \subseteq A$, and $A \cap B \subseteq B$

$\therefore$ $\overline{A \cap B} \subseteq \overline{A}$, and $\overline{A \cap B} \subseteq \overline{B}$. Hence $\overline{A \cap B} \subseteq \overline{A} \cap \overline{B}$.

The result can be extended to the intersection of an arbitrary family $\{A_\alpha\}$ of subsets of X, *i.e.*,

$$\overline{\bigcap_{\alpha \in \Lambda} A_\alpha} \subseteq \bigcap_{\alpha \in \Lambda} \overline{A_\alpha}$$

Remark: Let (X, d) be a metric space and $A \subseteq Y \subseteq X$. Then the closure of A in (Y, d_Y) is denoted by $\overline{A}^Y$. It is easy to verify that $\overline{A}^Y = \overline{A} \cap Y$.

2.8 Interior, Exterior, Frontier and Boundary Points

Definition. Let A be any subset of a metric space (X, d). A point 'a' in A is an *interior point* of A if there exists $r > 0$, such that

$$a \in S_r(a) \subseteq A.$$

Interior of a Set. The set of all interior points of A is called the *interior* of A, and we write

$$\text{int } A = \{x \in A : S_r(x) \subseteq A, \text{ for some } r > 0\}$$

clearly int A is an open subset of A.

A point $x \in X$ is said to be an *exterior point* of A, if it is an interior point of the complement of A. *i.e.*, if there exists an open sphere $S_r(x)$ such that

$$S_r(x) \subseteq A^c, \text{ or } S_r(x) \cap A = \phi.$$

Exterior of a Set. The set of all exterior points of A, denoted by ext A, is called the *exterior* of A.

By definition

$$\text{ext } A = \text{int } A^c$$

$\therefore$ ext A is open and is the largest open set contained in A^c.

Also a point is an *exterior point* of A if and only if it is not an adherent point of A.

$\therefore$ $$\text{ext } A = (\overline{A})^c$$

Moreover, int A = ext $A^c = (\overline{A^c})^c$, and int $A^c = (\overline{A})^c$.

For example, if $X = \mathbf{R}$, and $A = [0, 1[$, then

$$\text{int } A =]0, 1[$$

$$\text{ext } A =]-\infty, 0[\cup]1, \infty[.$$

Note that the points '0' and '1' are neither interior points nor exterior points of A. Such points of A are called *Frontier points* of A.

In another example, let A be the set of complex numbers $x + iy$, such that $y = -1$, and $0 \le x \le 1$. Then $A' = A$ and there are no interior points.

A point $x \in X$ is said to be a *frontier point* of $A \subseteq X$ if it is neither an interior nor an exterior point of A. If the frontier point belongs to A it is then called a *boundary point* of A.

Frontier and Boundary of a Set. The set of all frontier points and boundary points are denoted by $F_r(A)$, and $bd(A)$, respectively. Clearly

$$bd(A) \subseteq F_r(A).$$

In the above example $F_r(A) = \{0, 1\}$, and $bd(A) = \{0\}$. Note that the interior points, exterior points and frontier points of any subset A of X fill up the whole space X.

i.e.,
$$X = \text{int } A \cup \text{ext } A \cup F_r(A)$$

Now since int A and ext A both are open in X, therefore $F_r(A)$ is a closed set.

ILLUSTRATIONS

1. $X = \mathbf{R}$, and d the usual metric, $A = \mathbf{Q}$, then int $A = \phi$, ext $A = \phi$, $F_r(A) = \mathbf{R}$, and $bd(A) = \phi$.
2. $X = \mathbf{R}$, and d is the usual metric, $A =]1, 2] \cup]3, 4[$, then
$$\text{int } A = A, \text{ ext } A =]-\infty, 1[\cup]2, 3[\cup]4, \infty[$$
$$F_r(A) = \{1, 2, 3, 4\}, \text{ and } bd(A) = \{2\}.$$
3. If (X, d) is a discrete metric space, and $A \subseteq X$, then
$$\text{int } A = A, \text{ ext } A = A^c, F_r(A) = bd(A) = \phi.$$
4. If $X = \mathbf{N}$, and A be a finite subset of it, say $A = \{1, 2, 3\}$, then
$$\text{int } A = \phi, \text{ ext } A = \phi$$
$$F_r(A) = \mathbf{N}, \text{ and } bd(A) = A.$$

Properties:

Let A and B be any two subsets of a metric spaces (X, d), then

(*i*) int A is the largest open set contained in A.
i.e., $A = \cup\{G: G$ is open, and $G \subseteq A\}$

(*ii*) A is open if and only if $A = \text{int } A$.

(*iii*) $A \subseteq B$ implies int $A \subseteq$ int B

(*iv*) int $(A \cap B) = (\text{int } A) \cap \text{int } B$.

(*v*) int $(A \cup B) \supseteq \text{int } A \cup \text{int } B$.

Proof of (*i*) and (*ii*) follow from theorems 1 and 2, chapter 2 (change open interval to open sphere in the proof).

(*iii*) Let $x \in \text{int } A$. Then there exists $r > 0$ such that

$$S_r(x) \subseteq A$$

Therefore

$$S_r(x) \subseteq B \qquad (\because\ A \subseteq B)$$

This implies $x \in \text{int } B$.

Hence $\text{int } A \subseteq \text{int } B$.

(*iv*) By definition

$\text{int } A \subseteq A$, and $\text{int } B \subseteq B$

$\therefore \text{int } A \cap \text{int } B$, being the intersection of two open sets, is open. Therefore

$\text{int } A \cap \text{int } B \subseteq \text{int } (A \cap B)$ (by (*ii*))

Also $A \cap B \subseteq A \Rightarrow \text{int } (A \cap B) \subseteq \text{int } A$, and $\text{int } (A \cap B) \subseteq \text{int } B$

$\therefore\ \text{int } (A \cap B) \subseteq \text{int } A \cap \text{int } B$.

(*v*) $A \subseteq A \cup B \Rightarrow \text{int } A \subseteq \text{int } (A \cup B)$

and $B \subseteq A \cup B \Rightarrow \text{int } B \subseteq \text{int } (A \cup B)$

$\therefore\ \text{int } A \cup \text{int } B \subseteq \text{int } (A \cup B)$.

Note that equality may not hold in (*v*) as can be seen by the following example:

Let $A =]2, 5[$, and $B = [5, 7[$ be the subsets of the metric space $(\mathbf{R}, d)$ with the usual metric d, then $\text{int } A =]2, 5[$, $\text{int } B =]5, 7[$

$\text{int } A \cup \text{int } B =]2, 7[- \{5\}$

But $\text{int } (A \cup B) =]2, 7[$

$\therefore\ \text{int } (A \cup B) \neq \text{int } A \cup \text{int } B$

Theorem 7. *Let (X, d) be a metric space, and A, B be any two subsets of X, then*

(*i*) *ext A is the largest open set contained in A^c*

(*ii*) *A^c is open if and only if $A^c = ext\ A$*

(*iii*) *$A \subseteq B$ implies $ext\ B \subseteq ext\ A$*

(*iv*) *$ext\ (A \cap B) \supseteq ext\ A \cup ext\ B$*

(*v*) *$ext\ (A \cup B) = ext\ A \cap ext\ B$*

Proof follows from the above theorem by taking complements.

Theorem 8. Let *(X, d) be a metric space, and A, B are subset of X, then*

(*i*) $Fr(A) = \overline{A} \cap (\overline{A})^c = \overline{A} - int\ A$

(*ii*) *$Fr(A) = \phi$ if and only if A is both open and closed*

(*iii*) *A is closed if and only if $A \supseteq Fr(A)$*

(*iv*) *A is open if and only if $A^c \supseteq Fr(A)$*

(*v*) *$Fr(A \cap B) \subseteq Fr(A) \cup Fr(B)$. The equality holds if $\overline{A} \cap \overline{B} = \phi$*

(*vi*) *$Fr(int\ A) \subseteq Fr(A)$.*

(*i*) $Fr(A) = (\text{int } A \cup \text{ext } A)^c = (\text{int } A^c) \cap (\text{ext } A)^c = \overline{A}^c \cap \overline{A}$

i.e., $Fr(A) = \overline{A} \cap \overline{A}^c = \overline{A} - (\overline{A}^c)^c = \overline{A} - \text{int } A$

(*ii*) Let $Fr(A) = \phi$, then by (*i*)

$$\overline{A} - \text{int } A = \phi \quad i.e., \ \overline{A} \subseteq \text{int } A$$

$$\Rightarrow \qquad A \subseteq \overline{A} \subseteq \text{int } A \subseteq A$$

Hence, A is both open and closed.

Conversely, let A be both open and closed then

$$Fr(A) = \overline{A} - \text{int } A = A - A = \phi.$$

(*iii*) Let A be closed, then

$$Fr(A) = \overline{A} \cap (\overline{A}^c) \qquad \text{[by } (i)]$$

$$= A \cap (\overline{A}^c) \subseteq A$$

Conversely, let $Fr\ A \subseteq A$.

If possible, let A be not closed, then there exists an element x belonging to $\overline{A}$, but not belonging to A, *i.e.*, $x \in \overline{A} - A$.

But

$$\overline{A} - A = \overline{A} \cap A^c \subseteq \overline{A} \cap (\overline{A}^c) = Fr(A) \qquad \text{[by } (i)]$$

$$\therefore \qquad x \in Fr(A)$$

So that $x \in A$, which is a contradiction.

Hence, A must be closed.

(*iv*) A is open if and only if A^c is closed, and A^c is closed if and only if

$$A^c \supseteq Fr(A^c)$$

But since

$$Fr(A^c) = (\overline{A}^c) \cap \overline{(A^{c^c})} \qquad \text{[by } (i)]$$

$$= \overline{A}^c \cap \overline{A} = Fr(A)$$

Hence, A is open if and only if $A^c \supseteq Fr(A)$.

(*v*) $$Fr(A \cap B) = \overline{(A \cap B)} \cap \overline{(A \cap B)^c} \qquad \text{[by } (i)]$$

$$\subseteq \overline{A} \cap \overline{B} \cap \overline{(A^c \cup B^c)}$$

$$= \overline{A} \cap \overline{B} \cap (\overline{A}^c \cup \overline{B}^c)$$

$$= (\overline{A} \cap \overline{B} \cap \overline{A}^c) \cup (\overline{A} \cap \overline{B} \cap \overline{B}^c)$$

$$= (Fr(A) \cap \overline{B}) \cup (Fr(B) \cap \overline{A})$$

$$\subseteq (Fr(A) \cup Fr(B)$$

(*vi*) $$Fr(A \cup B) = Fr(A \cup B)^c = Fr(A^c \cap B^c) \subseteq Fr(A^c) \cup Fr(B^c)$$

$$= Fr(A) \cup Fr(B)$$

2.9 Dense Sets

Definition. A subset A of a metric space (X, d) is said to be *dense* (or *everywhere dense*) in X, if the closure of A is X, *i.e.*, $\overline{A} = X$.

For example the set of rationals is dense in **R** with the usual metric. Every interval is dense-in-itself.

A set A is said to be *nowhere dense* in X, if the complement of the closure of A is dense in X.

i.e.,
$$(\overline{A})^c = X, \text{ or } \overline{\text{ext } A} = X$$

Equivalently, A is nowhere dense in X iff int $(\overline{A}) = \phi$, since

$$\text{int } \overline{A} = \text{ext } (\overline{A})^c, \text{ and } \text{ext } X = \phi$$

Clearly every finite subset of X is nowhere dense.

A set is said to be *somewhere* dense in X, if it is not nowhere dense in X.

It is clear that A is somewhere dense in X if and only if $\overline{A}$ contains a non-empty open sphere.

Also since $\overline{A}$ is closed, A is nowhere dense in X if and only if $\overline{A}$ is nowhere dense in X. Clearly a subset of a nowhere dense set, is nowhere dense.

A set 'A' is said to be *dense-in-itself* if every point of A is a limit point of A.

i.e., $A \subseteq A'$

A set 'A' is said to be perfect if it is closed and dense-in-itself,

i.e., $A = A'$

Every closed interval, the empty set, ϕ, and the whole space X are perfect sets.

Example 17. The *cantor set* is a perfect set.

Recall that the cantor set is the set obtained from the closed interval $[0, 1]$ by removing the sequence of open intervals $\left]\frac{1}{3}, \frac{2}{3}\right[$; $\left]\frac{1}{9}, \frac{2}{9}\right[\cup \left]\frac{7}{9}, \frac{8}{9}\right[$; which are middle thirds of $[0, 1]$; $\left[0, \frac{1}{3}\right]$; $\left[\frac{2}{3}, 1\right]$;... respectively. Thus the *cantor set* is the intersection of the family of sets $\{F_n : n \in \mathrm{N}\}$, where

$$F_1 = [0, 1]$$

$$F_2 = \left[0, \frac{1}{3}\right] \cup \left[\frac{2}{3}, 1\right]$$

$$F_3 = \left[0, \frac{1}{9}\right] \cup \left[\frac{2}{9}, \frac{1}{3}\right] \cup \left[\frac{2}{3}, \frac{7}{9}\right] \cup \left[\frac{8}{9}, 1\right], \text{ etc.}$$

are closed sets.

($\because$ Each F_n is the complement of the union of removed open intervals and the intervals $]-\infty, 0[$ and $]1, \infty[$), and so the cantor set

$$F = \bigcap_{n=1}^{\infty} F_n$$

is closed.

All that remains is to show that it is dense-in-itself.

For this let $x \in F$, then $x = \sum_{k=1}^{\infty} \frac{a_k}{3^k}$, where each a_k is either 0 or 2, be the ternary expansion of x (expansion in the scale of 3). We shall show that x is a limit point of F.

Choose the sequence $\{x_n\}$ in F, such that

$$x_1 = \cdot a_1' a_2 a_3 \ldots a_n \ldots$$

$$x_2 = \cdot a_1 a_2' a_3 \ldots a_n \ldots$$

$$\vdots$$

$$x_n = \cdot a_1 a_2 \ldots a_n' \; a_{n+1} \ldots$$

$$\vdots$$

where $a_n' = 0$, if $a_n = 2$, and $a_n' = 2$, if $a_n = 0$.

The sequence $\{x_n\}$ of distinct points of F differ from x at the nth place in the ternary expansion.

Therefore $$\lim_{n \to \infty} x_n = x$$

Thus every point of the cantor set is a limit point of the set and so it is dense-in-itself.

Ex. Prove that the cantor set is nowhere dense.

Definition. A metric space X is said to be *separable* if there is a countable subset of X, which is dense in X.

Since the set of all rational numbers is countable, and dense in **R**, therefore the metric space **R** is separable. Let (X, d) be a discrete metric space where X is any uncountable set, then (X, d) is not separable ($\because$ the only dense subset of (X, d) is X itself). Hence the discrete metric space is separable if and only if it is countable.

Ex. 1. A subset A of (X, d) is dense in X if and only if A has non-empty intersection with each non-empty open sphere in X, or equivalently if and only if A has non-empty intersection with each non-empty open subset of X.

[*Hint*: Given $S_r(x) \cap A \neq \phi$, for all $r > 0$, and for each $x \in X$, and by definition

$$\overline{A} = \{x \in X : S_r(x) \cap A \neq \phi, \text{ for all } r > 0\}.$$

hence $\overline{A} = X$.]

Ex. 2. Show that the Euclidean space $\mathbf{R}^n$ is separable.

Ex. 3. Prove that the metric space l_∞ of all bounded sequences with sup metric is not separable.

[*Hint*: Show that every dense subset is uncountable.]

EXERCISE

1. Let x_1 and x_2 be distinct points in the metric space (X, d). Show that there exist two disjoint open spheres centered at x_1 and x_2, respectively.
2. Show by example that a set which fails to be closed need not be open.

3. Show by example that a non-empty proper subset A of a metric space (X, d) can be both open and closed.
4. Give an example to show that the closure of the open sphere of radius r centered at x_0 is not necessarily equal to the closed sphere or radius r centered at x_0. Prove that we always have
$$\overline{S}_r(x_0) \subseteq S_r[x_0]$$
5. Let $\{[a_n, b_n]\}$ be a sequence of closed intervals such that $|a_n| \leq 1, \ |b_n| \leq 1, \ \forall\ n$. Then prove that
$$\{\{x_n\} : x_n \in [a_n, b_n]\}$$
is a closed subset of H_∞.
6. If A and B are disjoint closed subsets of a metric space X, show that
$$G = \{x \in X : d(x, A) < d(x, B)\}, \text{ and } H = \{x \in X : d(x, B) < d(x, A)\}$$
are disjoint open sets containing A and B respectively.
7. Determine whether the following subsets of the metric spaces indicated are open, closed, both open and closed, or neither open nor closed
   (*i*) $\{(x_1, x_2 \ldots, x_n) \in \mathrm{R}^n : x_i > 0 \text{ for } i = 1, 2 \ldots n\}$ and d is the Euclidean metric.
   (*ii*) $\{(x_1, x_2 \ldots, x_n) \in \mathrm{R}^n : x_i \text{ is rational for } i = 1, 2 \ldots n\}$
   (*iii*) $\{\{x_n\} \in l : x_n < 1/n, \text{ for } n = 1, 2 \ldots\}$, where l is the set of all sequences $\{x_n\}$ such that $\sum_{n=1}^{\infty} |x_n|$ is convergent with the metric defined by
$$d(\{x_n\}, \{y_n\}) = \sum_{n=1}^{\infty} |x_n - y_n|.$$
   (*iv*) $\{f \in C'[a, b] : f(a) + f'(a) = 0\}$, where $C'[a, b]$ is the set of all functions defined on $[a, b]$ having continuous first order derivative on $[a, b]$, with the metric defined by
$$d(f, g) = \sup \{| f(x) - g(x) | : x \in [a, b]\}$$
$$+ \sup \{| f'(x) - g'(x) | : x \in [a, b]\}.$$
8. Give an example of a countable family of closed subsets of **R** whose union is not closed.
9. Prove that an open subset of $\mathbf{R}^n$ can be expressed as the union of a countable family of open spheres in $\mathbf{R}^n$.
10. Show that if $n \geq 2$, then there are open subsets of $\mathbf{R}^n$ which cannot be expressed as the union of a countable family of pairwise disjoint open spheres in $\mathbf{R}^n$.
11. Show that a metric space is discrete if and only if every point of the space is isolated.
12. Find the closures, the interiors, and the frontiers of the following:
   (*i*) a subset A of a discrete metric space X,
   (*ii*) the set of all rational numbers in **R**,
   (*iii*) an open sphere in the Euclidean space $\mathbf{R}^n$.
13. Prove that the cantor set of example 17 has neither isolated points nor interior points.
14. Let A be the subset
$$\left\{\left(\frac{m}{n}, \frac{1}{n}\right) : n = 1, 2, \ldots;\ m = 0, \pm 1, \pm 2, \ldots\right\}$$
of the Euclidean space $\mathbf{R}^2$. Prove that $\overline{A}$ is the union of A and the set $\{(x, 0) : x \in \mathbf{R}\}$.
15. Find the Frontier of the subset $\{x_1, x_2) : x_2 = 0\}$ of the Euclidean space $\mathbf{R}^2$.
16. Prove that a set D is dense in a metric space (X, d) if and only if X is the only closed set containing D.
17. Verify that $(\mathbf{R} - \mathbf{Q}, d)$, the irrationals with the usual metric, is a separable metric space.

18. Prove that every subspace of a separable metric space is separable.
19. Prove that the space $(\mathbf{R}^n, d)$ is separable, with the metric d given by $d(x, y) = \max_{1 \le k \le n} |x_k - y_k|$,

$$x = (x_1, x_2, \ldots, x_n), \text{ and } y = (y_1, \ldots, y_n) \in \mathbf{R}^n.$$

20. Prove that the example of a Hilbert space (Example 2) is a separable metric space.
21. Prove that a closed set in a metric space (X, d) either is nowhere dense in X or else contains some non-empty open set.
22. Prove that $A = \{f_n : n \in \mathbf{N}\}$ is a nowhere dense subset of $\mathbf{C}\,[0, 1]$ w.r.t. sup metric, where $f_n(x) = n - n^3x$, if $x \le 1/n^2$, and $f_n(x) = 0$, otherwise.

3. CONVERGENCE AND COMPLETENESS

Definition. Let (X, d) be any metric space. The sequence $\{a_n\}$ of points of X is said to *converge* to a point 'a' of X, if for each $\varepsilon > 0$ there exists a positive integer m, such that

$$d(a_n, a) < \varepsilon, \quad \forall\ n \ge m$$

i.e.,

$$d(a_n, a) \to 0, \quad \text{as } n \to \infty$$

or equivalently, for each open sphere $S_\varepsilon(a)$ centered at 'a' there exist a positive integer m such that a_n is in $S_\varepsilon(a)$, for all $n \ge m$.

The point 'a' is called the *limit* of the sequence $\{a_n\}$, and we write $a_n \to a$, as $n \to \infty$

i.e.,

$$\lim_{n\to\infty} a_n = a.$$

Cauchy sequence. A sequence $\{a_n\}$ of points of (X, d) is said to be a *Cauchy sequence* if for each $\varepsilon > 0$ there exists a positive integer n_0, such that

$$d(x_n, x_m) < \varepsilon, \quad \forall\ n, m \ge n_0$$

i.e.,

$$d(x_n, x_m) \to 0, \quad \text{as} \quad n, m \to \infty.$$

Theorem 9. *Every convergent sequence is a Cauchy sequence.*

Let (X, d) be any metric space. Let the sequence $\{a_n\}$ of points in X converge to a.

For every given $\varepsilon > 0$ there exists a positive integer n_0 such that

$$d(a_n, a) < \varepsilon/2, \quad \forall\ n \ge n_0$$

Then for $m, n \ge n_0$ we have

$$d(a_n, a_m) \le d(a_n, a) + d(a, a_m) < \varepsilon/2 + \varepsilon/2 = \varepsilon$$

This implies $\{a_n\}$ is a Cauchy sequence.

Note: The following examples show that the converse of the statement need not be true.

Example 18. Consider the space $X =]0, 1]$ of the real line with the usual metric. The sequence $\{a_n\} = \{1/n\}$ is a Cauchy sequence, converges to '0', which is not a point of the space.

Example 19. Let $\mathbf{Q}$ be the set of rational numbers in which the metric d is defined by

CHAPTER 19

$$d(x, y) = |x - y|, \quad \forall\, x, y \in \mathbf{Q}$$

($\mathbf{Q}$, d) is a metric space. The sequence $\{1/3^n\}$ is a Cauchy sequence which converges to the limit 0. But the sequence $\{(1 + 1/n)^n\}$ is also a Cauchy sequence in it, which does not converge to a point of $\mathbf{Q}$.

Complete metric space. A metric space (X, d) is said to be *complete* if every Cauchy sequence converges to a point of X.

The spaces in the examples mentioned above are not complete. But if we adjoin the point '0' to the space]0, 1] in the first example it becomes complete.

Remark: Any metric space which is not already complete can be made so by adjoining additional points to it.

ILLUSTRATIONS

1. The discrete space (X, d) is a complete metric space. For in this space a Cauchy sequence must be a constant sequence (*i.e.*, it must consist of a single point repeated from some place on) and so converges.
2. The space (R, d) is a complete metric space. The convergence in $\mathbf{R}$ is the ordinary convergence of numerical sequences.
3. The space R^n of all ordered n-tuples with the metric d,

$$d(x, y) = \left(\sum_{i=1}^{n} (x_i - y_i)^2 \right)^{1/2}$$

is a complete metric space. The convergence in this space is coordinate wise. This space $(\mathbf{R}^n, d)$ is called n-dimensional Euclidean space.

Example 20. The space C [0, 1] of all bounded continuous real-valued functions defined on the closed interval [0, 1] with the metric d given by

$$d(f, g) = \max_{0 \le x \le 1} |f(x) - g(x)|$$

is a complete metric space.

■ Let $\{f_n\}$ be a Cauchy sequence in C [0, 1].

Let $\varepsilon > 0$ be given. Then there exists a positive integer n_0, such that

$$d(f_n, f_m) < \varepsilon, \quad \forall\, n, m \ge n_0$$

i.e.,
$$\max_{0 \le x \le 1} |f_n(x) - f_m(x)| < \varepsilon, \quad \forall\, n, m \ge n_0$$

i.e.,
$$|f_n(x) - f_m(x)| < \varepsilon, \quad \forall\, n, m \ge n_0 \text{ and } \forall\, x \in [0, 1]$$

By Cauchy criterion of uniform convergence, the sequence of functions $\{f_n\}$ converges uniformly on [0, 1]. Let f be the limit of a uniformly convergent sequence of continuous functions so this itself is continuous on [0, 1]. Hence the Cauchy sequence $\{f_n\}$ converges to a point of C [0,1].

Example 21. Let l_∞ be the set of all bounded numerical sequences $\{x_n\}$ in which the metric d is defined by

$$d(x, y) = \sup_n |x_n - y_n|, \forall x = \{x_n\}, y = \{y_n\} \in l_\infty$$

■ Let $\{x_n\}$ be a Cauchy sequence of elements of l_∞ and let

$$x_n = \{a_i^{(n)}\}. \text{ Since } x_n \in l_\infty, \text{ so } \exists M > 0,$$
$$|a_i^{(n)}| \le M, \text{ for } i = 1, 2, 3...$$

Therefore for $\varepsilon > 0$, there exists an integer n_0 such that

$$d(x_n, x_m) < \varepsilon, \quad \forall n, m \ge n_0$$

i.e.,
$$\sup_i |a_i^{(n)} - a_i^{(m)}| < \varepsilon, \quad \forall n, m \ge n_0$$

$\Rightarrow$
$$|a_i^{(n)} - a_i^{(m)}| < \varepsilon, \quad \forall n, m \ge n_0, \text{ and for all } i = 1, 2, 3... \qquad ...(1)$$

Let i be fixed. Then (1) implies that the sequence $\{a_i^{(1)}\ a_i^{(2)},...\ a_i^{(n)},...\}$ is Cauchy and so converges to a_i, by Cauchy's General Principle of convergence. Taking limit in (1) as $m \to \infty$, we have

$$|a_i^{(n)} - a_i| \le \varepsilon, \quad \forall n \ge n_0$$

and this is true for all $i = 1, 2, 3,...$

Hence,
$$|a_i| \le |a_i^{(n)} - a_i| + |a_i^{(n)}| < \varepsilon + M, \forall i$$

This implies $\{a_i\}$ is bounded. Let $x = \{a_i\}$. Then $x \in l_\infty$. Hence (l_∞, d) is a complete space.

Example 22. Let l_p be the set of all real numerical sequences for which

$$\sum_{i=1}^{\infty} |x_i|^p < \infty.$$

■ We define the metric d in l_p by

$$d(x, y) = \left(\sum_{i=1}^{\infty} |x_i - y_i|^p\right)^{1/p}, \quad \forall x = \{x_i\}, y = \{y_i\} \in l_p$$

The space (l_p, d) is a complete metric space, and is known as Hilbert sequence space.

Consider a Cauchy sequence $\{x_n\} = \{\{x_i^{(n)}\}\}$ in l_p.

Therefore for a given $\varepsilon > 0$ there exists an integer n_0, such that

$$d(x_n, y_m) = \left(\sum_{i=1}^{\infty} |x_i^{(n)} - x_i^{(m)}|^p\right)^{1/p} < \varepsilon, \quad \forall n, m \ge n_0 \qquad ...(1)$$

Hence $|x_i^{(n)} - x_i^{(m)}| < \varepsilon$, $\forall n, m \ge n_0$, and for all $i \in \mathrm{N}$ ...(2)

Fixing i, we see that the sequence $\{x_i^{(1)}\ x_i^{(2)},...\ x_i^{(n)},...\}$ converges to a limit x_i

i.e.,
$$\lim_{n \to \infty} x_i^{(n)} = x_i$$

Let $x = \{x_i\}$. Then the inequality (1) implies

$$\sum_{i=1}^{k} | x_i^{(n)} - x_i^{(m)} |^p < \varepsilon^p, \text{ for every } k, \text{ and for } n, m \geq n_0$$

Taking limit as $m \to \infty$, we have

$$\sum_{i=1}^{k} | x_i^{(n)} - x_i |^p \leq \varepsilon^p, \quad \text{for} \quad n \geq n_0$$

Letting $k \to \infty$, we get

$$\sum_{i=1}^{\infty} | x_i^{(n)} - x_i |^p \leq \varepsilon^p, \quad \text{for} \quad n \geq n_0$$

This implies $x_n - x \in l_p$, and so $x = x_n - (x_n - x) \in l_p$. Also $d(x_n, x) \to 0$ as $n \to \infty$. Hence l_p is complete.

Example 23. Let X be the set of all continuous real-valued functions defined on [0, 1], and let

$$d(x, y) = \int_0^1 | x(t) - y(t) | \, dt, \quad x, y \in X$$

Show that (X, d) is not complete.

■ Let $\{x_n\}$ be a sequence in X defined by

$$x_n(t) = \begin{cases} n, & \text{if } 0 \leq t \leq \dfrac{1}{n^2} \\ \dfrac{1}{\sqrt{t}}, & \text{if } \dfrac{1}{n^2} \leq t \leq 1 \end{cases}$$

For $n > m$, we have

$$\begin{aligned} d(x_n, x_m) &= \int_0^1 | x_n(t) - x_m(t) | \, dt \\ &= \int_0^{1/n^2} | n - m | \, dt + \int_{1/n^2}^{1/m^2} \left| \frac{1}{\sqrt{t}} - m \right| dt + \int_{1/m^2}^{1} \left| \frac{1}{\sqrt{t}} - \frac{1}{\sqrt{t}} \right| dt \\ &= \frac{(n-m)}{n^2} + (2t^{1/2} - mt) \Big|_{1/n^2}^{1/m^2} \\ &= \frac{1}{n} - \frac{m}{n^2} + \left(\frac{2}{m} - \frac{1}{m} \right) - \left(\frac{2}{n} - \frac{m}{n^2} \right) \\ &= \frac{1}{m} - \frac{1}{n} \to 0, \text{ as } m, n \to \infty. \end{aligned}$$

Hence $\{x_n\}$ is a Cauchy sequence in X.

Now we shall show that this Cauchy sequence does not converge in X. For every $x \in X$

$$d(x_n, x) = \int_0^1 |x_n(t) - x(t)|\, dt$$

$$= \int_0^{1/n^2} |n - x(t)|\, dt + \int_{1/n^2}^1 \left| \frac{1}{\sqrt{t}} - x(t) \right| dt$$

Since integrals are non-negative, so is each integral on the right, and hence $d(x_n, x) \to 0$, as $n \to \infty$ would imply that each integral approaches zero, and since x is in X, so x is continuous.

But

$$x(t) = \begin{cases} t^{-1/2}, & \text{if } 0 < t \le 1 \\ 0\ , & \text{if } t = 0 \end{cases}$$

which is discontinuous at $t = 0$. Hence $d(x_n, x)$ does not tend to zero for each $x \in X$, *i.e.*, the sequence $\{x_n\}$ does not converge to the point of the space. This implies that (X, d) is not complete.

Lemma. *Let (X, d) be any metric space and A be any non-empty subset of X, then $x \in \overline{A}$ if and only if there exists a sequence $\{x_n\}$ in A such that $x_n \to x$, as $n \to \infty$.*

Let $x \in \overline{A}$ then every open sphere centered at x intersects A. In particular $S_{1/n}(x) \cap A \neq \phi$, for all n. So we get a sequence $\{x_n\}$ in A such that

$$d(x_n, x) < \frac{1}{n}, \forall\, n$$

$$\Rightarrow \qquad \lim_{n\to\infty} x_n = x.$$

Again, let $\{x_n\}$ be a sequence in A which converges to x. To show that $x \in \overline{A}$ we must show that every open sphere centered at x intersects A.

Let $S_r(x)$ be any open sphere. Then for $r > 0$, $\lim_{n\to\infty} x_n = x$ implies that there exists a positive integer n_0, such that

$$d(x_n, x) < r, \quad \forall\, n \ge n_0$$

In particular

$$d(x_{n_0}, x) < r$$

$$\Rightarrow \qquad x_{n_0} \in S_r(x)$$

$$\Rightarrow \qquad S_r(x) \cap A \neq \phi \qquad [\because x_{n_0} \in A]$$

$$\Rightarrow \qquad x \in \overline{A}$$

Theorem 10. *Let (X, d) be a complete metric space and Y be a subspace of X. Then Y is complete if and only if it is closed in (X, d).*

Let Y be a complete subspace of X. In order to show that Y is closed we need to show that $Y = \overline{Y}$. By definition $Y \subset \overline{Y}$, so we shall show that $\overline{Y} \subseteq Y$.

Let x be an element of Y. If $x \in \overline{Y}$, the result is proved. If $x \notin Y$, then x is a limit point of Y. By definition of limit point, every neighbourhood $S_{1/n}(x)$ contains at least one member of Y other than x. Thus for each n we get a sequence $\{y_n\}$ in Y such that

$$d(y_n, x) < 1/n. \text{ Thus } y_n \to y, \text{ as } n \to \infty.$$

Now the sequence $\{y_n\}$ being a convergent sequence must be a Cauchy sequence. Since Y is complete, this Cauchy sequence $\{y_n\}$ must converge in Y, hence $x \in Y$. But x is an arbitrary point of $\overline{Y}$, therefore $\overline{Y} \subseteq Y$.

Conversely, we assume that Y is a closed subspace of X, and establish that Y is complete.

Let $\{y_n\}$ be a Cauchy sequence in Y, and since X is given to be a complete space, therefore $\{y_n\}$ must converge to a point y in X. But then $y_n \in Y$, for all n, and $y_n \to y$, as $n \to \infty$

$$\Rightarrow \qquad y \in \overline{Y} = Y \qquad (\because Y \text{ is closed})$$

The following is a generalisation of Nested-Interval Theorem (Example 17, Chapter 3).

Theorem 11. ***Cantor's Intersection Theorem.*** *Let (X, d) be a complete metric space, and let $\{F_n\}$ be a decreasing sequence of non-empty closed subsets of X such that $d(F_n) \to 0$, as $n \to \infty$. Then $F = \bigcap_{n=1}^{\infty} F_n$ contains exactly one point.*

Since $F_n \neq \phi$, for each $n \in \mathbf{N}$, we can choose a sequence of points $\{x_n\}$ such that $x_n \in F_n$, for $n = 1, 2, 3...$ We shall show that $\{x_n\}$ is a Cauchy sequence in X.

Now $\{F_n\}$ is a decreasing sequence, *i.e.*, $F_{n+1} \subseteq F_n$, for all n, therefore $x_n, x_{n+1}, \ldots$ all lie in F_n.

Moreover $d(F_n) \to 0$, as $n \to \infty$. Therefore given $\varepsilon > 0$ there exists a positive integer n_0, such that

$$d(F_n) < \varepsilon, \quad \forall\, n \geq n_0$$

$$\Rightarrow \qquad x_{n_0}, x_{n_{0+1}}, x_{n_{0+2}}, \ldots \text{ all lie in } F_{n_0}$$

Thus for positive integers $m, n \geq n_0$, we have

$$d(x_n, x_m) \leq d(F_{n_0}) < \varepsilon$$

$\Rightarrow$ $\{x_n\}$ is a Cauchy sequence in X. Since (X, d) is complete, $\exists$ a point $x \in X$ such that $\lim_{n \to \infty} x_n = x$.

We shall now show that $x \in \bigcap_{n=1}^{\infty} F_n$

If possible, $x \notin \bigcap_{n=1}^{\infty} F_n$. Then there exists a positive integer m, such that $x \notin F_m$. Since F_m is closed, and $x \notin F_m$.

$$d(x, F_m) > 0. \text{ Let } d(x, F_m) = r > 0,$$

then
$$d(x, y) \geq r, \quad \forall\, y \in F_m.$$

Thus, the open sphere $S_{r/2}(x)$, and F_m are clearly disjoint, and therefore

$$n > m \Rightarrow F_n \subset F_m,$$

and this implies

$$x_n \in F_m \ (\because x_n \in F_n) \Rightarrow x_n \notin S_{r/2}(x).$$

This is impossible, since $\{x_n\}$ converges to x. Hence $x \in \bigcap_{n=1}^{\infty} F_n$.

Now to show that $x \in \bigcap_{n=1}^{\infty} F_n$ is unique.

If possible, let y be another point in $\bigcap_{n=1}^{\infty} F_n$

Then $y \in F_n$, for every n.

$\Rightarrow \quad d(x, y) \leq d(F_n)$, for every n (by the definition of the diameter).

But, since it is given that $d(F_n) \to 0$ as $n \to \infty$.

Therefore on taking limit as $n \to \infty$.

$d(x, y) \leq 0$. But $d(x, y) \geq 0$ is always true. Hence $d(x, y) = 0$, and so $x = y$.

Note: The converse of the above theorem is also true.

Ex. If every decreasing sequence of non-empty closed sets whose diameter tends to zero have a non-empty intersection in a metric space (X, d), then (X, d) is complete.

The following examples show that the condition $\lim_{n \to \infty} d(F_n) = 0$, and that the sets F_n's are closed, are both necessary in the above theorem.

Example 24. Let X be the real line **R** with the usual metric, and let $F_n = [n, \infty[$.

Now $X = \mathbf{R}$ is complete. The sets F_n are closed and $F_1 \supset F_2 \supset F_3 \ldots \supset F_n \ldots$ But $\bigcap_{n=1}^{\infty} F_n$ is empty. Observe that $\lim_{n \to \infty} d(F_n) \neq 0$.

Example 25. Let X be the real line with the usual metric, and let $F_n =]0, 1/n]$

Now $X = \mathbf{R}$ is complete

$$F_1 \supset F_2 \supset F_3 \ldots\ldots \supset F_n$$

and
$$\lim_{n \to \infty} d(F_n) = 0, \text{ but } \bigcap_{n=1}^{\infty} F_n \text{ is empty.}$$

Observe that $F_n's$ are not closed.

Definition. A subset A of a metric space (X, d), possibly the whole space, is said to be of the *first category*, if it is the union of a countable family of nowhere dense sets.

i.e., $A \subseteq X$ can be written as $A = \bigcup_{n=1}^{\infty} A_n$,

where each A_n is nowhere dense in X, *i.e.*, int $(\overline{A}_n) = \phi$, for each n.

Otherwise it is said to be of the *second category*. It is important to note, that in a discrete space the only nowhere dense set is the empty set, *i.e.*, every non-empty set is of the second category. In particular the set I of integers is of the second category in the space **R** of reals with the discrete metric. On the other hand if $(\mathbf{R}, d)$ is a metric space with the usual metric, then the set **I** of integers is nowhere dense and hence is of the first category. This shows that the set is not of the first or second category in and on its own, rather its category classification also depends on the metric space to which the set belongs.

Ex. Prove that

(*i*) Q is of first category in **R**, w.r.t. usual metric,

(*ii*) every countable subspace of **R** is of first category in **R**,

(*iii*) if X is of second category, and if $X = A \cup B$, then either A or B must be of second category,

(*iv*) X is of second category in itself if and only if the intersection of every countable family of dense open sets in X is non-empty,

(*v*) if A is a dense subset of a complete metric space X, and if $A = \bigcap_{n=1}^{\infty} G_n$, where $G_n's$ are open in X, then $X - A$ is of first category.

Theorem 12. ***Baire's Category Theorem.*** *If $\{A_n\}$ is a sequence of nowhere dense sets in a complete metric space (X, d), then*

$$X \neq \bigcup_{n=1}^{\infty} A_n.$$

i.e., Every complete metric space is of second category.

To prove the theorem we need the following lemma.

Lemma. *Let A be a nowhere dense subset of a metric space, (X, d). Let G be any non-empty open set in X,* and *$r > 0$ be any real number, then there exists an open sphere of radius less than or equal to r contained in G and disjoint from A.*

Since A is nowhere dense *i.e.*, int $\overline{A} = \phi$, and int $\overline{A}$ is the largest open set containing $\overline{A}$. Therefore, if G is any non-empty open set, then

$$G \not\subseteq \overline{A},$$

G being non-empty and $G \not\subseteq \overline{A}$, therefore $\exists$ an $x \in G$ such that $x \notin \overline{A}$.

Moreover G is open, $\exists$ an open sphere $S_r(x)$, for some $r > 0$, such that

$$S_r(x) \subset G.$$

Since $x \notin \overline{A}$, we can choose a positive number $r_1 < r$, such that

$$S_{r_1}(x) \subset S_r(x), \text{ and } S_{r_1}(x) \cap A = \phi.$$

Thus $$x \in S_{r_1}(x) \subset G, \text{ and } S_{r_1}(x) \cap A = \phi.$$

Proof of the main theorem:

X is open, being a non-empty open subset of itself, and A is nowhere dense in X, then by the above lemma, for given $r_1 > 0, 0 < r_1 < 1$, $\exists$ an open sphere $S_{r_1}(x_1)$ in X, such that

$$S_{r_1}(x_1) \cap A_1 = \phi$$

Let F_1 be the concentric closed sphere of radius $\frac{1}{2}r_1$,

i.e., $$F_1 = S_{\frac{1}{2}r_1}[x_1],$$

and consider its interior, int $F_1 \neq \phi$. Let $x_2 \in$ int F_1. Since A_2 is nowhere dense, int F_1 contains an open sphere $S_{r_2}(x_2)$ of radius $r_2 < \frac{1}{2}$, such that

$$S_{r_2}(x_2) \cap A_2 = \phi$$

Let F_2 be the concentric closed sphere of radius $\frac{1}{2}r_2$, *i.e.*, $F_2 = S_{r_2/2}[x_2]$. Since A_3 is nowhere dense, int F_2, being non-empty, contains an open sphere $S_{r_3}(x_3)$ centered at $x_3 \in$ int F_2 and radius $r_3 < \frac{1}{3}$, such that

$$S_{r_3}(x_3) \cap A_3 = \phi$$

Let F_3 be the concentric closed sphere of radius $\frac{1}{2}r_3$.

Continuing in this manner we get a decreasing sequence $\{F_n\}$ of non-empty closed subsets of X, where

$$F_n = S_{r_n/2}[x_n].,$$

and

$$F_{n+1} = S_{\frac{1}{2}r_{n+1}}[x_{n+1}] \subset S_{r_{n+1}}(x_{n+1}) \subseteq S_{\frac{1}{2}r_n}[x_n] = F_n$$

i.e., $$F_{n+1} \subseteq F_n, \quad \forall\, n.$$

Also

$$d(F_n) = 2\, r_n/2 < 1/n, \quad \forall\, n$$

$\therefore$ $$d(F_n) \to 0, \text{ as } n \to \infty$$

Since X is given to be complete, therefore by Cantor's intersection theorem we conclude that $\bigcap_{n=1}^{\infty} F_n$ contains exactly one point say $x \in X$.

$\Rightarrow$ $$x \in F_n, \quad \forall\, n$$

$\Rightarrow$ $$x \in S_{\frac{1}{2}r_n}[x_n] \subset S_{r_n}(x_n), \quad \forall\, n$$

and $$S_{r_n}(x_n) \cap A_n = \phi$$

$$\Rightarrow \qquad x \notin A_n \quad \forall\, n$$

$$\Rightarrow \qquad x \notin \bigcup_{n=1}^{\infty} A_n$$

Hence,
$$\bigcup_{n=1}^{\infty} A_n \neq X.$$

Ex. Use Baire's Category theorem to prove the existence of everywhere continuous, nowhere differentiable real-valued functions.

[*Hint*: Take

$$A_n = \{f \in C[0,1] : \exists\, x \in [0, 1 - 1/n], \text{ such that}$$
$$\left|\frac{f(x+h) - f(x)}{h}\right| \le n, \text{ for } 0 < h < 1/n\}.$$

If $f \in C[0,1]$ has a derivative at some point, then $f \in A_n$, for some n. Show that A_n is closed, and has empty interior.]

4. CONTINUITY AND UNIFORM CONTINUITY

Definition. Let (X, d_1), and (Y, d_2) be any two metric spaces. A function $f : X \to Y$ is said to be *continuous* at a point 'a' of X, if for given $\varepsilon > 0$, there exists a $\delta > 0$, such that

$$d_2(f(x), f(a)) < \varepsilon, \text{ whenever } d_1(x, a) < \delta.$$

Equivalently, for each open sphere $S_\varepsilon(f(a))$ centered at $f(a)$ there is an open sphere $S_\delta(a)$ centered at a such that

$$f(S_\delta(a)) \subseteq S_\varepsilon(f(a))$$

The function $f : X \to Y$ is said to be continuous, if it is continuous at each point of X.

Example 26. If (X, d) is a discrete space then every function $f : X \to Y$ is continuous on X.

■ For any $a \in X$ if we choose $\delta < 1$. Then

$$S_\delta(a) = \{a\}$$

and so
$$f(S_\delta(a)) = \{f(a)\} \subseteq S_\varepsilon(f(a)) \text{ holds for each positive } \varepsilon.$$

Example 27. If (X, d_1), and (Y, d_2) are any two metric spaces; then the constant function $f : X \to Y$ is continuous on X.

Theorem 13. *Let (X, d_1), and (Y, d_2) be any two metric spaces and f is a function from X into Y. Then f is continuous at $a \in X$ if and only if, for every sequence $\{a_n\}$ converging to 'a' we have*

$$\lim_{n\to\infty} f(a_n) = f(a)$$

i.e.,
$$a_n \to a \Rightarrow f(a_n) \to f(a).$$

First, let us suppose that the function f is continuous at a point $a \in X$ and $\{a_n\}$ is a sequence in X, such that $\lim_{n\to\infty} a_n = a$.

Since f is continuous at a, therefore for any given $\varepsilon > 0$, there exists a $\delta > 0$, such that

$$d_2(f(x), f(a)) < \varepsilon, \text{ whenever } d_1(x, a) < \delta \quad ...(1)$$

Again, since $\lim_{n\to\infty} a_n = a$, therefore $\exists$ a positive integer n_0 such that

$$d_1(a_n, a) < \delta, \quad \forall\, n \geq m$$

From (1) putting $x = a_n$, we have

$$d_2(f(a_n), f(a)) < \varepsilon, \text{ whenever } d_1(a_n, a) < \delta, \quad \forall\, n \geq m$$

$\Rightarrow$ $\quad d_2(f(a_n), f(a)) < \varepsilon, \quad \forall\, n \geq m$

$\Rightarrow$ $\quad \{f(a_n)\}$ converges to $f(a)$

i.e., $\quad \lim_{n\to\infty} f(a_n) = f(a)$.

We now assume that f is not continuous at a; and show that though there exists a sequence $\{a_n\}$ converging to 'a' yet the sequence $\{f(a_n)\}$ does not converge to $f(a)$.

Since f is not continuous at a, therefore there exists at least one $\varepsilon > 0$ such that for every $\delta > 0$ $d_2(f(x), f(a)) \geq \varepsilon$, and $d_1(x, a) < \delta$, for some $x \in X$.

Therefore, by taking $\delta = \frac{1}{n}$, we find that for each positive integer n there is $a_n \in X$ such that

$$d_2(f(a_n), f(a)) \geq \varepsilon, \text{ and } d_1(a_n, a) < \frac{1}{n}$$

Thus, the sequence $\{f(a_n)\}$ does not converge to $f(a)$.

This shows that continuous functions of one metric space into another are those functions which send every convergent sequence to a convergent sequence or in other words which preserve convergence.

Theorem 14. *Let (X, d_1), and (Y, d_2) be two metric spaces, then $f : X \to Y$ is continuous if and only if $f^{-1}(G)$ is open in X, whenever G is open in Y.*

We first assume that f is continuous. If G is any open subset in Y, we shall show that $f^{-1}(G)$ is open in X. (Recall that if G is a subset of Y then the set $\{x \in X : f(x) \in G\}$, consists of all points of X whose images lie in G, is denoted by $f^{-1}(G)$). If $f^{-1}(G) = \phi$, it is open; so we assume that $f^{-1}(G) \neq \phi$. Let $x \in f^{-1}(G)$. Then $f(x) \in G$ and since G is open, $\exists$ an open sphere $S_\varepsilon(f(x))$ such that

$$S_\varepsilon(f(x)) \subseteq G, \text{ for some } \varepsilon > 0$$

Now by definition of continuity, there exists an open sphere $S_\delta(x)$ such that

$$f(S_\delta(x)) \subseteq S_\varepsilon(f(x)), \text{ for } \delta > 0$$

But $S_\varepsilon(f(x)) \subseteq G$.

$\therefore \quad S_\delta(x) \subseteq f^{-1}(G) \Rightarrow f^{-1}(G)$ is open

CHAPTER 19

Now we assume that $f^{-1}(G)$ is open in X, whenever G is open in Y, and show that f is continuous. Let x be an arbitrary point in X, and let $\varepsilon > 0$ be given. Let $S_\varepsilon(f(x))$ be an open sphere in Y centered at $f(x)$.

This open sphere in an open set, so its inverse image is an open set which contains x.

i.e., $$f^{-1}(S_\varepsilon(f(x))) \text{ is open in } X$$

Since $x \in f^{-1}(S_\varepsilon(f(x)))$, $\exists$ a $\delta > 0$ such that

$$S_\delta(x) \subseteq f^{-1}(S_\varepsilon(f(x)))$$

This implies $f(S_\delta(x)) \subseteq S_\varepsilon(f(x))$

Hence, f is continuous as x. Since x was taken to be an arbitrary point of X. Hence f is continuous at every point of X.

Note: From the above theorem we observe that continuous functions are precisely those which pull open sets back to open sets. It is to be noted that a continuous function need not take open sets to open sets.

Example 28. Let $f : \mathbf{R} \to \mathbf{R}$ be defined by

$$f(x) = \sin x$$

in the metric space $(\mathbf{R}, d)$. It can be easily seen that f is continuous. The open set $]0, 2\pi[$ in $(\mathbf{R}, d)$ is mapped to the closed set $[-1, 1]$ in $(\mathbf{R}, d)$.

Example 29. Let $f : \mathbf{R} \to \mathbf{R}$ be defined by

$$f(x) = x^2.$$

f is continuous on $\mathbf{R}$. f maps the open set $]-2, 2[$ onto the semi-closed set $[0, 4[$.

Ex. The function $f : X \to Y$ of a metric space (X, d_1) into a metric space (Y, d_2) is continuous iff the inverse image of every closed set contained in Y is closed.

[*Hint*: This follows from the preceding theorem by taking complements.]

Example 30. Let (X, d_1) and (Y, d_2) be metric spaces. Show that $f : X \to Y$ is continuous if and only if $f(\overline{A}) \subseteq \overline{f(A)}$, for every $A \subseteq X$.

■ Let f be continuous and A be any subset of X. Then $f^{-1}(\overline{f(A)})$ is closed in X.

Now $$f(A) \subseteq \overline{f(A)}$$

$$\Rightarrow \quad A \subseteq f^{-1}(\overline{f(A)})$$

$$\Rightarrow \quad \overline{A} \subseteq \overline{f^{-1}(\overline{f(A)})} = f^{-1}(\overline{f(A)})$$

i.e., $$f(\overline{A}) \subseteq \overline{f(A)}.$$

Conversely, suppose $$f(\overline{A}) \subseteq \overline{f(A)}, \qquad \ldots(1)$$

for every subset A of X.

Let F be any closed subset of Y, then

$$f^{-1}(F) \subseteq X$$

using (1), we have

$$f(\overline{f^{-1}(F)}) \subseteq \overline{f(f^{-1}(F))} = \overline{F} = F$$

$$\therefore \quad \overline{f^{-1}(F)} \subseteq f^{-1}(F)$$

Thus $f^{-1}(F)$ is closed in X.

Hence f is continuous.

Definition. Let (X, d_1) and (Y, d_2) be two metric spaces. A function $f : X \to Y$ is said to be ***uniformly continuous*** if for each $\varepsilon > 0$ there exists a $\delta > 0$ (depending on ε alone) such that

$$d_2(f(x), f(y)) < \varepsilon, \text{ whenever } d_1(x, y) < \delta, \quad \forall\, x, y \in X$$

Every function $f : X \to Y$ which is uniformly continuous on X is necessarily continuous on X.

However, converse may not be true. We shall see later that these two concepts are equivalent on concept metric spaces.

Example 31. For any non-empty subset A of a metric space (X, d) the function $f : X \to \mathbf{R}$ given by

$$f(x) = d(x, A), \text{ for } x \in X$$

is uniformly continuous. Also show that $f(x) = 0 \Leftrightarrow x \in \overline{A}$.

■ By definition we have for $x \in X$

$$d(x, A) = \inf\{d(x, a) : a \in A\}$$

By triangle inequality,

$$d(x, a) \le d(x, y) + d(y, a), \ \forall\, a \in A \subseteq X, \ x, y \in X$$

On taking infimum

$$d(x, A) = \inf_{a \in A} d(x, a) \le d(x, y) + \inf_{a \in A} d(y, a)$$

[$\because$ $d(x, y)$ is independent of a]

$$= d(x, y) + d(y, A)$$

$$\therefore \quad d(x, A) - d(y, A) \le d(x, y) \quad \text{...(1)}$$

Since (1) is true for each $x, y \in X$. Therefore on interchanging x and y,

$$d(y, A) - d(x, A) \le d(y, x) = d(x, y).$$

Thus

$$|d(x, A) - d(y, A)| \le d(x, y) \quad \text{...(2)}$$

Now for each $\varepsilon > 0$, choose a $\delta \le \varepsilon$

then

$$|f(x) - f(y)| = |d(x, A) - d(y, A)| \le d(x, y) < \delta \le \varepsilon \qquad \text{[using (2)]}$$

i.e., $\quad |d(x, A) - d(y, A)| < \varepsilon,$ whenever $d(x, y) < \delta$

Hence f is uniformly continuous on X.

For the second part, let $f(x) = 0$, *i.e.*, $d(x, A) = 0$.

This implies that there exists a sequence $\{a_n\}$ in A such that

$$d(a_n, x) \to d(x, A)$$

i.e.,
$$\lim_{n\to\infty} d(a_n, x) = d(x, A) = 0$$

$\Rightarrow$
$$a_n \to x, \text{ as } n \to \infty$$

Therefore for a given $\varepsilon > 0$, there exists a positive integer n_0, such that

$$a_n \in S_\varepsilon(x), \qquad \forall\, n \ge n_0$$

In particular $a_{n_0} \in S_\varepsilon(x)$. But $a_{n_0} \in A$.

Therefore for each $\varepsilon > 0$, $S_\varepsilon(x)$ contains a point of A other than x.

Hence $x \in \overline{A}$.

Conversely, let $x \in \overline{A}$, then there exists a sequence $\{x_n\}$ in A such that $\{x_n\}$ converges to x.

$\therefore$
$$d(x_n, x) \to 0 \text{ as } n \to \infty$$

But, since $d(x, A) \le d(x, x_n)$, $\forall\ n \in \mathrm{N},\ x_n \in A$

and
$$d(x, x_n) \to 0, \text{ as } n \to \infty$$

$\therefore\ d(x, A) \le 0$. Hence $d(x, A) = 0$

Example 32. Let (X, d) be a metric space then show that any disjoint pair of closed sets in X can be separated by disjoint open sets in X.

■ Let A and B be any closed subsets of X such that

$$A \cap B = \phi.$$

Define a mapping $f : X \to \mathbf{R}$ by

$$f(x) = \frac{d(x, A)}{d(x, A) + d(x, B)}, \quad x \in X$$

f is well-defined.

Since $d(x, A) + d(x, B) \ne 0$, $\forall x \in X$. For if $d(x, A) + d(x, B) = 0$ for some $x \in X$, then $d(x, A) = 0$ and $d(x, B) = 0$.

This implies $x \in \overline{A} = A$ and $x \in \overline{B} = B$

i.e., $\quad x \in A \cap B$, which is impossible $(\because A \cap B = \phi)$

f is continuous on X, since the functions $x \to d(x, A)$, and $x \to d(x, B)$ are continuous on X.

Clearly $$f(x) = \begin{cases} 0, & \text{if} \quad x \in A \\ 1, & \text{if} \quad x \in B \end{cases}$$

Let $$G = \{x \in X : f(x) < \tfrac{1}{2}\}$$

then $G = f^{-1}(]-\infty, \frac{1}{2}[)$, being an inverse image of an open interval $]-\infty, \frac{1}{2}[$ under a continuous mapping f is an open subset of X.

Moreover $x \in A$ implies $f(x) = 0 < \frac{1}{2}$, *i.e.*, $x \in G$

$\therefore$ $$A \subseteq G.$$

Similarly, $H = \{x \in X : f(x) > \frac{1}{2}\}$ is an open set of X containing B. Also $G \cap H = \phi$. Hence the result.

Definition. Let (X, d), and (Y, d') be any two metric spaces. A function $f: X \to Y$ is said to be a **homeomorphism** if

(*i*) f is both one-one and onto,

(*ii*) f and f^{-1} are both continuous.

By theorem 14 it follows that a homeomorphism induces a 1 – 1 correspondence between the open sets in X and open sets in Y.

Two metric spaces are said to be *homeomorphic* if there exists a homeomorphism between them. But not all metric properties are shared by homeomorphic spaces as is shown by the following example.

Example 33. Let $X = \{1, \frac{1}{2}, \frac{1}{3}, ...,\}$ with d the usual metric on the subsets of **R**, and let $Y = \mathbf{N}$, the set of natural numbers with the usual metric $d'(= d)$. Then the function $f : X \to Y$ defined by $f(1/n) = n$ is a homeomorphism from (X, d) to (Y, a'). In fact every subset of X and every subset of Y is open in the respective spaces. But (X, d) is a bounded metric space and (Y, d') is not. Also (Y, d') is a complete metric space but (X, d) is not.

Definition. A function $f: X \to Y$ is called an *isometry* if

$$d(x, y) = d'(f(x), f(y)), \quad \forall\ x, y \in X.$$

Clearly each isometry is always one-to-one and uniformly continuous.

Two metric spaces are said to be *isometric* if there exists an isometry between them which is onto. It is easy to verify that if two metric spaces are isometric, then they are necessarily homeomorphic. But its converse may not be true as can be seen by the above example.

By definition it follows that isometric spaces possess all the same metric properties. Such spaces are metrically identical and differ only in names of their elements.

Theorem 15. *The image of a Cauchy sequence under a uniformly continuous function is again a Cauchy sequence.*

Let (X, d_1), and (Y, d_2) be two metric spaces and $f : X \to Y$ be uniformly continuous. Let $\{x_n\}$ be any Cauchy sequence in X; and let $\varepsilon > 0$ be given. Then, f being uniformly continuous, there exists a $\delta > 0$ (depending on ε) such that

$$d_2(f(x_m), f(x_n)) < \varepsilon, \text{ whenever } d_1(x_m, x_n) < \delta \quad \ldots(1)$$

Since $\{x_n\}$ is Cauchy, corresponding to this $\delta > 0$ there exists a positive integer n_0 (depending on δ and so on ε) such that

$$d_1(x_m, x_n) < \delta, \text{ for } m, n \geq n_0$$

From (1) and (2), we conclude that

$$d_2(f(x_m), f(x_n)) < \varepsilon, \text{ for } m, n \geq n_0$$

Hence $\{f(x_n)\}$ is a Cauchy sequence in Y.

Remark: Note that a continuous function may not send a Cauchy sequence to a Cauchy sequence; as can be seen by the following example.

Let X be the set of all positive real numbers and d the usual metric on X, and $Y = \mathbf{R}$ be the set of all real numbers with d' the usual metric on $\mathbf{R}$. Then $f : X \to Y$ defined by

$$f(x) = \frac{1}{x}, \forall\, x \in X, \text{ is continuous on } X$$

Now $\left\{\frac{1}{n}\right\}$ is a Cauchy sequence in X $\left[\because \left|\frac{1}{n} - \frac{1}{m}\right| \to 0 \text{ as } m, n \to \infty\right.$. But $\left\{f\left(\frac{1}{n}\right)\right\} = \{n\}$ is not a Cauchy sequence in Y. Hence f cannot be a uniformly continuous function.

Theorem 16. *Let (X, d_1) be a metric space, and (Y, d_2) be a complete metric space. If f is a uniformly continuous function from a subset A of X into Y, then f can be extended uniquely to a uniformly continuous function g from A into Y.*

We shall prove the theorem in the following steps.

(1) Existence of $g : \overline{A} \to Y$

(2) Uniform continuity of g

(3) Uniqueness of g.

(1) Let $\{a_n\}$ be any convergent sequence in A converging to a point $x \in \overline{A}$. Also $\{a_n\}$ being convergent must be a Cauchy sequence and since f is uniformly continuous, its image $\{f(a_n)\}$ is a Cauchy sequence in Y. Again Y is given to be complete, the sequence $\{f(a_n)\}$ must be a convergent sequence in Y, and so there exists a point y in Y such that

$$f(a_n) \to y, \text{ i.e., } \lim_{n \to \infty} f(a_n) = y$$

Now we shall show that y depends only on x and not on the sequence $\{a_n\}$.

Let $\{b_n\}$ be another sequence in A such that $\{b_n\}$ converges to x then by the triangle inequality in (X, d_1) we have

$$d_1(a_n, b_n) \leq d_1(a_n, x) + d_1(x, b_n)$$

$$\Rightarrow \quad d_1(a_n, b_n) \to 0, \text{ as } n \to \infty \quad (\because a_n \to x, \text{ and } b_n \to x, \text{ as } n \to \infty)$$

And by the uniform continuity of f

$$d_2(f(a_n), f(b_n)) \to 0, \text{ as } n \to \infty$$

Now from the triangle inequality in (X, d_2)

$$d_2(f(b_n), y) \le d_2(f(b_n), f(a_n)) + d_2(f(a_n), y)$$

we have

$$d_2(f(b_n), y) \to 0, \text{ as } n \to \infty$$

i.e.,
$$\lim_{n\to\infty} f(b_n) = y.$$

This shows that y is independent of the sequence $\{a_n\}$ in A.

Thus if we define

$$y = g(x)$$

Then g extends f from A to $\overline{A}$ which can be seen as follows:

Let $x \in A$, then $x \in \overline{A}$.

Taking $a_n = x, \forall\ n$, the sequence $\{a_n\}$ is a constant sequence in A and so $a_n \to x$.

Then $g(x) = \lim_{n\to\infty} f(a_n)$

But since $f(a_n) = f(x)$ we get

$$\lim_{n\to\infty} f(a_n) = f(x), \forall\ x \in A$$

$\therefore\ f(x) = g(x), \forall\ x \in A$

Thus g extends f to $\overline{A}$.

(2) Let $\varepsilon > 0$ be given. By uniform continuity of f we can find $\delta > 0$ such that for all $a, b \in A$ we have

$$d_1(a, b) < \delta \quad \Rightarrow \quad d_2(f(a), f(b)) < \varepsilon \qquad \text{...(1)}$$

Let x, y be any point in $\overline{A}$ such that

$$d_1(x, y) < \delta$$

there exist sequences $\{a_n\}$ and $\{b_n\}$ in A such that $a_n \to x$ and $b_n \to y$ respectively.

i.e.,
$$d_1(a_n, x) \to 0 \text{ as } n \to \infty, \text{ and } d_1(b_n, y) \to 0 \text{ as } n \to \infty$$

For $r = \dfrac{\delta - d_1(x, y)}{2} > 0$, $\exists$ a positive integer n_0 (depending on r) such that

$$d_1(a_n, x) < r, d_1(b_n, y) < r, \quad \forall\ n \ge n_0$$

Now
$$d_1(a_n, b_n) \le d_1(a_n, x) + d_1(x, y) + d_1(y, b_n)$$
$$< r + d_1(x, y) + r = \delta, \quad \forall\ n \ge n_0$$

It follows from (1) that

$$d_2(f(a_n), f(b_n)) < \varepsilon, \quad \forall\ n \ge n_0 \qquad \text{...(2)}$$

By definition of g, we have

$$f(a_n) \to g(x), \text{ and } f(b_n) \to g(y) \text{ as } n \to \infty$$

i.e., for each $\varepsilon > 0, \exists$ positive integers m_1, m_2 such that

$$d_2(f(a_n), g(x)) < \varepsilon/3, \quad \forall\, n \geq m_1$$

and

$$d_2(f(b_n), g(x)) < \varepsilon/3, \quad \forall\, n \geq m_2$$

By triangle inequality in Y

$$d_2(g(x), g(y)) \leq d_2(g(x), f(a_n)) + d_2(f(a_n), f(b_n)) + d_2(f(b_n), g(y)).$$

Using (2) and definition of g,

$$d_2(g(x), g(y)) < \varepsilon/3 + \varepsilon/3 + \varepsilon/3 = \varepsilon, \quad \forall\, n \geq m = \max\,(m_1, m_2, n_0)$$

Thus $\quad d_1(x, y) < \delta \Rightarrow d_2(g(x), g(y)) < \varepsilon, \quad \forall\, x, y \in \overline{A}$

Hence g is uniformly continuous.

(3) We shall now show that g is unique. Let, if possible, there be another extension $h : \overline{A} \to Y$ of f to $\overline{A}$ such that h is uniformly continuous.

We have for all $x \in A$

$$g(x) = f(x) = h(x)$$

and for all $x \in \overline{A}$

$$g(x) = f(x) = h(x), \text{ by taking limits}$$

Hence $\quad g(x) = h(x), \quad \forall\, x \in \overline{A}$

This shows that g is unique.

4.1 Banach Fixed Point Theorem

Definition. Let (X, d) be a metric space. A mapping $f: X \to X$ is said to be a *contraction mapping*, if there exists a positive real number α with $\alpha < 1$, such that

$$d(f(x), f(y)) \leq \alpha\, d(x, y), \quad \forall\, x, y \in X$$

we observe that, applying f to each of the two points of the space contracts the distance between them. Obviously f is continuous.

Example 34. If $x = \{x_n\} \in l_2$, then $f(x) = \left\{\dfrac{x_n}{2}\right\}$ is a contraction mapping on l_2. For if $y = \{y_n\}$ is any other point of l_2, then

$$d(f(x), f(y)) = \left(\sum_{n=1}^{\infty} \left(\frac{x_n}{2} - \frac{y_n}{2} \right)^2 \right)^{1/2} = \frac{1}{2}\, d(x, y)$$

Example 35. If $f(x) = x^2, 0 \leq x \leq \dfrac{1}{3}$. Then f is a contraction mapping on $[0, \frac{1}{3}]$ with the usual metric d.

$$d(f(x), f(y)) = d(x^2, y^2) = |x^2 - y^2| = |x - y||x + y| < \frac{2}{3}|x - y|$$

$$\therefore \quad d(f(x), f(y) \le \frac{2}{3} d(x, y)$$

Definition. A point $x \in X$ is called a *fixed point* of the mapping $f : X \to X$, if $f(x) = x$.

Theorem 17. ***Banach fixed point theorem.*** *Any contraction mapping f of a non-empty complete metric space (X, d) into itself has a unique fixed point.*

For all $x, y \in X$, we have $d(f(x), f(y)) \le \alpha d(x, y)$ ...(1)

for some $\alpha, 0 < \alpha < 1$

This implies that f is continuous.

Now choose any point $x_0 \in X$. Let us define a sequence $\{x_n\}$ by

$$x_1 = f(x_0), x_2 = f(x_1), \ldots, x_{n+1} = f(x_n), \ldots$$

Then

$$x_n = f^n(x_0), \quad \forall\, n \in \mathbf{N}$$

We shall show that the sequence $\{x_n\}$ is Cauchy. For each positive integer n we have

$$\begin{aligned} d(x_n, x_{n+1}) &= d(f(x_{n-1}), f(x_n)) \le \alpha d(x_{n-1}, x_n) \\ &\le \alpha^2 d(x_{n-2}, x_{n-1}) \\ &\le \alpha^3 d(x_{n-3}, x_{n-2}) \\ &\vdots \\ &\le \alpha^n d(x_0, x_1) \end{aligned} \qquad \text{...(2)}$$

By triangle inequality, we have for $n \ge m$

$$\begin{aligned} d(x_m, x_n) &\le d(x_m, x_{m+1}) + d(x_{m+1}, x_{m+2}) + \ldots + d(x_{n-1}, x_n) \\ &\le \alpha^m d(x_0, x_1) + \alpha^{m-1} d(x_0, x_1) + \ldots + \alpha^{n-1} d(x_0, x_1) \\ &= \alpha^m [1 + \alpha + \alpha^2 + \ldots + \alpha^{n-m-1}] d(x_0, x_1) \\ &= \frac{\alpha^m (1 - \alpha^{n-m})}{1 - \alpha} d(x_0, x_1) \\ &< \frac{\alpha^m d(x_0, x_1)}{1 - \alpha} \to 0, \text{ as } m \to \infty \qquad [\because \alpha < 1] \end{aligned}$$

Hence $\{x_n\}$ is a Cauchy sequence, and X being complete implies $x_n \to x$, for some $x \in X$.

Since f is continuous, therefore we have

$$f(x) = f(\lim_{n\to\infty} x_n) = \lim_{n\to\infty} f(x_n) = \lim_{n\to\infty} x_{n+1} = x$$

Hence,

$$f(x) = x$$

CHAPTER 19

To prove uniqueness, suppose $f(y) = y$, for some $y \in X$ then

$$d(x, y) = d(f(x), f(y)) \leq \alpha\, d(x, y)$$

Since $\alpha < 1$, and $d(x, y) \geq 0$ therefore we must have

$$d(x, y) = 0, \quad i.e., x = y.$$

EXERCISE

1. Let $\{x_n^k\}$ be a sequence in H_∞ (H_∞ as defined in Q. 9 Exercise, p. 716). Let $x = \{x_n\} \in H_\infty$, then prove that $\{x^k\}$ converges to x in H_∞ if and only if $\lim\limits_{k\to\infty} x_n^k = x_n, \ \forall\, n \in \mathrm{N}$

   [*Hint*: Here $d(\{x_n\}, \{y_n\}) = \sum\limits_{n=1}^{\infty} \dfrac{|x_n - y_n|}{2^n}$, $|x_n| \leq 1$, $|y_n| \leq 1$, $\forall\, n \in \mathbf{N}$, suppose $\lim\limits_{k\to\infty} x_n^k = x_n$, $\forall\, n$.

   Choose a positive integer m such that $\sum\limits_{n=m+1}^{\infty} \dfrac{2}{2^n} < \varepsilon/2$. Show that $\exists$ a positive integer m_1 such that if

   $$k \geq m_1, \text{ then } \sum_{n=m+1}^{\infty} \frac{|x_n^k - x_n|}{2^n} < \varepsilon/2$$

   Deduce that if $k \geq m_1$

   $$d(x^k, x) < \varepsilon$$

   For the converse note that

   $$d(x^k, x) = \sum_{n=1}^{\infty} \frac{|x_n^k - x_n|}{2^n} < \sum_{n=1}^{\infty} \varepsilon/2^n, \quad \forall\, n \geq m_1]$$

2. Let $\lim\limits_{n\to\infty} x_n = x$, and $\lim\limits_{n\to\infty} y_n = y$, where $\{x_n\}$ and $\{y_n\}$ are sequences of real numbers in the metric space (X, d) and $x, y \in X$. Prove that the sequence of real numbers $\{d(x_n, y_n)\}$ converges to the real number $d(x, y)$.

   [*Hint*: $d(x_n, y_n) \leq d(x_n, x) + d(x, y) + d(y, y_n)$].

3. Show that a Cauchy sequence is convergent $\Leftrightarrow$ It has a convergent subsequence.

4. Prove that if (X, d) is a complete space, and each $x \in X$ is a limit point of X, then X is uncountable.

5. Give an example of a complete metric space (X, d) and a sequence of non-empty closed sets $\{A_n\}$ in X with $A_1 \supseteq A_2 \supseteq A_3 \ldots \supseteq A_n \ldots$ such that $\bigcap\limits_{n=1}^{\infty} A_n = \phi$

   [*Hint*: The space **R** of real numbers with the usual metric is a complete metric space. Consider the sequence of closed sets

   $$A_n = [n, \infty[, n \in \mathbf{N}, \quad \bigcap_{n=1}^{\infty} A_n = \phi]$$

6. Give an example of a homeomorphism $f: X \to Y$ and a Cauchy sequence $\{x_n\}$ in X for which $\{f(x_n)\}$ is not Cauchy in Y.

   [*Hint*: Let $X = \{1, \frac{1}{2}, \frac{1}{3} \ldots\}$ with d the usual metric and d' the discrete metric. Then the identity function $I : (X, d) \to (X, d')$ is a homeomorphism, but the Cauchy sequence $\{1/n\}$ in (X, d) has its image $\{1/n\}$ in (X, d') which is not Cauchy].

7. Let X and Y be metric spaces and A be a non-empty subset of X. If f and g are continuous function from X into Y such that $f(x) = g(x)$, for every x in A, show that $f(x) = g(x)$ for every $x \in \overline{A}$.
8. Let f be a continuous real-valued function defined on $\mathbf{R}$ which satisfies the functional equation $f(x+y) = f(x) + f(y)$. Show that the function must have the form $f(x) = mx$ for some real number m.
9. Let $\{x_n\} \in l_2$. Prove that f defined by $f\{x_n\} = \sum_{n=1}^{\infty} a_n x_n$, a_n's are real numbers, is a continuous real-valued function on l_2.
10. Let f be a real-valued function on a metric space X. Prove that f is continuous on X if and only if the following sets $\{x : f(x) < c\}$ and $\{x : f(x) > c\}$ are open in X for every $c \in \mathbf{R}$.
11. Let (X, d) be a metric space such that $d(x, y) \leq 1$, $\forall\, x, y \in X$, and let $\{x_n\}$ be a sequence in X. If $f(x) = \{d(x, x_n)\}$, $x \in X$, then prove that
    (*i*) f is a continuous function from X into H_∞.
    (*ii*) If the range set of $\{x_n\}$ is dense in X, then f is one to one.
12. Give an example of a function $f : X \to Y$ which is one-to-one onto Y and continuous on X but not a homeomorphism. [*Hint*: Let $X = [0, 1]$ and let d be the usual metric and d' the discrete metric. Then the identity function $I : (X, d') \to (X, d)$ is one to one continuous but not a homeomorphism.]
13. If (X, d) is a complete metric space and if $\mathbf{F}$ is a family of real-valued continuous functions defined on X such that the set $\{f(x) : f \in \mathbf{F}\}$ is bounded for every $x \in X$. Then there is a non-empty open set $G \subseteq X$, and an $M > 0$ such that $|f(x)| \leq M$ for every $x \in G$ and for every $f \in \mathbf{F}$. [This is known as *uniform boundedness principle*].
14. Prove that if $f : X \to Y$ is an isometry from X onto Y, then for every Cauchy sequence $\{x_n\}$ in X, $\{f(x_n)\}$ is a Cauchy sequence in Y.
15. Prove that if the spaces (X, d) and (Y, d') are isometric then either they are both complete or neither is complete.
16. Let (X, d) be a metric space with $x_0 \in X$. Define $f : X \to \mathbf{R}$ by $f(x) = d(x, x_0)$. Prove that f is uniformly continuous on X.
17. Let $\phi : C[0,1] \to C[0,1]$ be defined by $\phi(f) = \alpha \int_0^x t\, f(t)\,dt$, where α is a constant.
    (*a*) Find α such that ϕ is a contraction mapping.
    (*b*) Also show that for each value of α, ϕ has a unique fixed point.
18. Prove the 'converse' of the Banach's fixed point theorem: if for each non-empty closed subset A of a metric space X, and for each contraction mapping $f : A \to A$, f has a fixed point, then X is complete.

CHAPTER 19

5. COMPACTNESS

The concept of compactness is an abstraction of an important property known as 'Heine Borel Property' posed by subsets of $\mathbf{R}$ which are closed and bounded. Heine Borel theorem states that if $I \subseteq \mathbf{R}$ is a closed interval, any family of open interval in R whose union contains I has a finite subfamily whose Compactness is concerned with covering the sets by open sets. Before defining compactness we need the following definitions.

Definition. Let (X, d) be a metric space. A family of subsets $\{A_\alpha\}$ in X is called a *cover* of any subset A of X if $A \subseteq \bigcup_{\alpha \in \Lambda} A_\alpha A$, Λ is any non-empty index set. If each A_α, $\alpha \in \Lambda$ is open in X, then the cover $\{A_\alpha\}$ is called an *open cover* of A.

A subfamily of the family $\{A_\alpha\}$ which itself is an open cover is called an *open subcover* of A. If the number of members in the subfamily is finite it is called a *finite subcover* of A.

Definition. A subset A of a metric space (X, d) is said to be *compact* if every open cover of A admits of a finite subcover, *i.e.*, for each family of open subsets $\{G_\alpha\}$ of X for which $\bigcup_{\alpha \in \Lambda} G_\alpha \supseteq A$, there exists a finite subfamily say $\{G_{\alpha_1}, G_{\alpha_2}, \ldots, G_{\alpha_n}\}$ such that $A \subseteq \bigcup_{i=1}^{n} G_{\alpha_i}$.

A metric space (X, d) is *compact* if X is itself compact, *i.e.*, for each family of open subsets $\{G_\alpha\}$ of X for which $\bigcup_{\alpha \in \Lambda} G_\alpha = X$, there exists a finite subfamily $\{G_{\alpha_i} : i = 1, 2, \ldots n\}$ such that

$$X = \bigcup_{i=1}^{n} G_{\alpha_i}$$

ILLUSTRATIONS

1. Any closed interval with the usual metric is compact.
2. The discrete space (X, d), where X is a finite set, is compact.
3. The space $(\mathbf{R}, d)$ where $\mathbf{R}$ is the set of reals and d is the usual metric is not compact, for the cover $\{]-n, n[: n \in \mathbf{N}\}$ is such that $\bigcup_{n=1}^{\infty}]-n, n[= \mathbf{R}$, which do not have a finite subcover.

Example 36. Prove that the open interval $]0, 1[$ with the usual metric is not compact.

■ The family of open intervals $\left\{\left]\frac{1}{n}, 1\right[: n = 2, 3\ldots\right\}$ is such that $\bigcup_{n=2}^{\infty} \left]\frac{1}{n}, 1\right[=]0, 1[$. Therefore $\left\{\left]\frac{1}{n}, 1\right[: n = 2, 3, \ldots\right\}$ is an open cover of $]0, 1[$, which has no finite subcover..

Example 37. Let X be an infinite set with the discrete metric. Show that (X, d) is not compact.

■ For each $x \in X$, $\{x\}$ is open in X

Also $\bigcup_{x \in X} \{x\} = X$

Therefore $\{\{x\} : x \in X\}$ is an open cover of X and since X is infinite, this open cover has no finite subcover.

Theorem 18. *Every closed subset of a compact metric space is compact.*

Let (X, d) be any compact metric space and F be any non-empty closed subset of X. We shall show that F is compact.

Let $\{G_\alpha : \alpha \in \Lambda\}$ be a family of open sets in X such that

$$\bigcup_{\alpha \in \Lambda} G_\alpha \supseteq F$$

Then $(\bigcup_{\alpha \in \Lambda} G_\alpha) \cup (X - F)$ is an open cover of X and by compactness of X, it has a finite subcover, say,

$$G_{\alpha_1}, G_{\alpha_2} \ldots G_{\alpha_n}, X - F$$

$\therefore$
$$\bigcup_{i=1}^{n} G_{\alpha_i} \cup (X - F) = X$$

or
$$\bigcup_{i=1}^{n} G_{\alpha_i} \supseteq F$$

Hence F is compact.

Note: This theorem shows that each closed subset of a compact metric space is compact. On the real line the closed set **N** is not compact in **R**. Also the closed set $F =]0, 1[$ is not compact in (X, d) where $X =]0, 2]$ and d is the usual metric. Observe that in these examples the space is not compact.

Theorem 19. *Every compact subset F of a metric space (X, d) is closed.*

Let F be any compact set. To prove that F is closed we shall show that F^c is open.

Let $y \in F^c$ and $x \in F$ then $x \neq y$

$\therefore$ $d(x, y) > 0$; let $d(x, y) = r_x$. Then the open spheres $S_{\frac{1}{2}r_x}(x)$ and $S_{\frac{1}{2}r_x}(y)$ are such that

$$S_{\frac{1}{2}r_x}(x) \cap S_{\frac{1}{2}r_x}(y) = \phi$$

For if z belongs to both $S_{\frac{1}{2}r_x}(x)$ and $S_{\frac{1}{2}r_x}(y)$, then

$$d(z, x) < \frac{1}{2}r_x \quad \text{and} \quad d(z, y) < \frac{1}{2}r_x$$

and by the triangle inequality

$$d(z, x) \leq d(x, z) + d(z, y) < \frac{1}{2}r_x + \frac{1}{2}r_x = r_x$$

which contradicts the fact that $d(x, y) = r_x$.

Now consider the collection $\{S_{\frac{1}{2}r_x}(x) : x \in F\}$ of open spheres of F. This collection is such that

$$\bigcup_{x \in F} S_{\frac{1}{2}r_x}(x) \supseteq F$$

Since F is compact, there exists a finite number of open spheres, say,

$$S_{\frac{1}{2}r_{x_1}}(x_1), S_{\frac{1}{2}r_{x_2}}(x_2), \text{K}, S_{\frac{1}{2}r_{x_n}}(x_n)$$

such that

$$\bigcup_{i=1}^{n} S_{r_{x_i}/2}(x_i) \supseteq F$$

Let $A_y = \bigcap_{i=1}^{n} S_{\frac{1}{2}r_{x_i}}(y)$. The set A_y is an open set, being the intersection of open spheres, containing y.

Since $S_{\frac{1}{2}r_{x_i}}(x) \cap S_{\frac{1}{2}r_{x_i}}(y) = \phi$, for each i, therefore we have

$$S_{\frac{1}{2}r_{x_i}}(x) \cap A_y = \phi$$

And so $(\bigcup_{i=1}^{n} S_{\frac{1}{2}r_{x_i}}(x_i)) \cap A_y = \phi$

This implies

$$F \cap A_y = \phi, \text{ or } A_y \subseteq F^c$$

Now $\bigcup_{y \in F^c} A_y = F^c$ and each A_y is open, therefore F^c is open and hence F is closed.

Note: The converse of the above theorem may not be true. For example $[0, \infty[$, is closed but not compact. The family of intervals $\{[0, n[: n \in \mathbf{N}\}$ is such that each set $[0, n[$ is open in $[0, \infty[$ and

$$\bigcup_{n=1}^{\infty} [0, n[= [0, \infty[.$$

This family is an open cover of $[0, \infty[$, which has no finite subcover.

Moreover, if a set in any metric space is not closed then it cannot be compact. This can be seen as follows:

Let A be any non-closed set in any metric space (X, d) and let a be a limit point of A which is not in A. Then the family of sets

$$(X - S_{1/n}(a) : n = 1, 2, 3, \ldots\}$$

is an open cover of A for which there is no finite subcover. Hence A is not compact.

Corollary. *A subset A of a compact metric space (X, d) is itself compact if and only if it is closed in (X, d).*

Theorem 20. *Every compact subset A of a metric space (X, d) is bounded.*

Suppose that A is compact and consider an open cover of A consisting of open spheres of radii–1

i.e.,
$$A \subseteq \bigcup_{x \in A} S_1(x)$$

Since A is compact, there must exist $x_1, x_2 \ldots x_n$ such that

$$A \subseteq \bigcup_{i=1}^{n} S_1(x_i)$$

Now let $M = \max d(x_i, x_j), \quad 1 \le i \le j \le n.$

Let $x, y \in A$ be any two elements then there exist elements x_i and x_j such that

$$x \in S_1(x_i), \text{ and } y \in S_1(x_j)$$

$\therefore$ By triangle inequality

$$d(x, y) \le d(x, x_i) + d(x_i, x_j) + d(x_j, y) \le 1 + M + 1 = M + 2$$

Hence A is bounded.

Note: Every compact subset of a metric space is closed and bounded but the converse need not be true, which can be seen by the following counter example. Let $X = [0, 1]$, and suppose (X, d) is a discrete space. Then the set $A = \{1, \frac{1}{2}, \frac{1}{3}, \ldots\}$ is closed because every point of A is a limit point of A, it is also bounded because $d(x, y) < 1$, for $x, y \in A$ and so $d(A) \le 1$. But the open cover $\left\{S_1\left(\frac{1}{n}\right)\right\} = \left\{\frac{1}{n}\right\}$ does not admit of a finite subcover. Hence A is not compact.

Theorem 21. *Heine-Borel theorem. Every closed and bounded subset of the real line is compact.*

Let A be any closed and bounded subset of $\mathbf{R}$. Since A is bounded there exist real numbers a and b such that $A \subseteq [a, b]$.

A being a closed subset of $\mathbf{R}$, it is a closed subset of $[a, b]$. Now since every closed subset of a compact metric space is compact therefore it is enough to show that $[a, b]$ is compact.

If $a = b$, there is nothing to prove, so we assume $a < b$. If possible, let $[a, b]$ be not compact, then there exists an open cover $\{G_\alpha\}$ of $[a, b]$ which has no finite subcover.

Let $a_1 = a,\ b_1 = b$ and $c_1 = \dfrac{a_1 + b_1}{2}$

So $[a, b] = [a_1, b_1] = [a_1, c_1] \cup [c_1, b_1]$.

At least one of these intervals in the union cannot be covered by a finite subfamily of $\{G_\alpha\}$. Let $[a_2, b_2]$ denote one of these intervals with this property. Thus $[a_2, b_2] \subseteq [a_1, b_1]$, and $b_2 - a_2 = \dfrac{1}{2}(b_1 - a_1)$,

Let $c_2 = \dfrac{a_2 + b_2}{2}$.

Therefore $[a_2, b_2] = [a_2, c_2] \cup [c_2, b_2]$. As before at least one of these intervals in the union cannot be covered by a finite subfamily of $\{G_\alpha\}$. Denote that interval by $[a_3, b_3]$. Then,

$$[a_3, b_3] \subseteq [a_2\ b_2] \subseteq [a, b_1], \text{ and } b_3 - a_3 = \frac{1}{2}(b_2 - a_2) = \frac{1}{2^2}(b_1 - a_1)$$

Continuing this process, we obtain a sequence $\{[a_n, b_n]\}$ of closed intervals such that each interval $[a_n, b_n]$ cannot be covered by a finite subfamily of $\{G_\alpha\}$, and

$$[a_{n+1}, b_{n+1}] \subseteq [a_n, b_n], \quad \forall\, n \in \mathbf{N}$$

with $b_n - a_n = \dfrac{1}{2^{n-1}}(b_1 - a_1)$ which tends to zero as n tends to ∞.

i.e., $$d[a_n, b_n] = |\, a_n - b_n \,| \to 0, \text{ as } n \to \infty.$$

Since $\mathbf{R}$ is complete, and $\{[a_n, b_n]\}$ is a non-empty decreasing sequence of closed subsets of $\mathbf{R}$ such that

$$d([a_n, b_n]) \to 0, \text{ as } n \to \infty$$

Therefore, by Cantor's intersection theorem

$\bigcap_{n=1}^{\infty} [a_n, b_n]$ contains exactly one element say x.

i.e., $$x \in [a_n, b_n], \ \forall\, n, \text{ and } [a_n, b_n] \subseteq [a_1, b_1] \subseteq [a, b], \forall\, n \qquad \ldots(1)$$

And so $x \in [a, b]$. Now since $\{G_\alpha\}$ is an open cover of $[a, b]$, therefore $x \in G_\alpha$ for some α, where

G_α is open in $[a, b]$.

This implies

$$G_\alpha = H \cap [a, b], \text{ where } H \text{ is open in } \mathbf{R}$$

Now $x \in G_\alpha \Rightarrow x \in H$

But H is open in $\mathbf{R}$. Therefore there exists an open sphere

$$S_\varepsilon(x) =]\, x - \varepsilon, x + \varepsilon\, [, \varepsilon > 0$$

such that

$$x \in]\, x - \varepsilon, x + \varepsilon\, [\subseteq H \qquad ...(2)$$

Since $b_n - a_n = \dfrac{1}{2^{n-1}}(b_1 - a_1) \to 0$, as $n \to \infty$

Therefore, for above $\varepsilon > 0$, there exists a positive integer n_0 such that

$$b_n - a_n < \varepsilon, \ \forall\, n \geq n_0$$

In particular,

$$b_{n_0} - a_{n_0} < \varepsilon$$

Therefore $[a_{n_0}, b_{n_0}] \subseteq]\, x - \varepsilon, x + \varepsilon\, [\subseteq H$, by (1) and (2)

Now $[a_{n_0}, b_{n_0}] \subseteq [a, b]$, and $[a_{n_0}, b_{n_0}] \subseteq H$

Therefore $[a_{n_0}, b_{n_0}] \subseteq G_\alpha$, *i.e.*, $[a_{n_0}, b_{n_0}]$ is covered by a single set G_α of the cover $\{G_\alpha\}$ which contradicts the fact that $[a_{n_0}, b_{n_0}]$ is not covered by any finite subfamily of $\{G_\alpha\}$. Hence our assumption that $[a, b]$ is not compact is wrong. Thus $[a, b]$ is compact.

Note: Converse of Heine-Borel theorem is also true. If A is a compact subset of $\mathbf{R}$, then A is closed and bounded. Let A be a compact subset of $\mathbf{R}$, and $\bigcup_{n=1}^{\infty}]-n, n[= \mathbf{R} \supseteq A$. So there exists positive integers $n_1, n_2, \ldots, n_k$ such that $\bigcup_{i=1}^{k}]-n_i, n_i[\supseteq A$. Take
$m = \max(n_1, n_2, \ldots, n_k)$, then $A \subseteq]-m, m[$.

Example 38. Consider the bounded set $A =]0, 1]$, A is not closed. Since 0 is a limit point of A which does not belong to A. Let $G = \{]1/n, 2[\ ; n \in \mathbf{N}\}$. G is an open cover of A and there is no finite subset of G which is a cover of A.

Example 39. Let $A = [0, \infty]$, A is a closed set, but it is not bounded. Consider the family of the sets

$$G = \{]n - 2, n\,[: n \in \mathbf{N}\}.$$

G is an open even cover of A. But G has no finite subcover for A.

Remarks:

1. Heine-Borel theorem does not hold in a general metric space as can be seen by the following example:
Let (X, d) be the discrete space and X is infinite, X is bounded as the distance between any two of its members is at the most one, X being the whole space is closed. Also for $\varepsilon = \frac{1}{2}$,

$$S_\varepsilon(x) = \{x\},\ x \in X \text{ so } X = \bigcup_{x \in X} S_\varepsilon(x),$$

Thus, the open cover $\{S_\varepsilon(x) : x \in X\}$ of X admits of no finite subcover.

Hence X is not compact.

2. However, from Theorems 19 and 20 it follows that the converse of Heine-Borel theorem is true for any general metric space.

Theorem 22. *Continuous image of a compact set is compact.*

Let (X, d_1) be a compact metric space and f be a continuous function from X into the metric space (Y, d_2) then $f(X)$, the image of X under f is compact in Y. Let $\{V_\alpha\}$ be any open cover of $f(X)$; which we denote by $Y_1 (Y_1 \subseteq Y)$, *i.e.*, each V_α is open in Y_1, and

$$Y_1 = \bigcup_{\alpha \in \Lambda} V_\alpha,\ \Lambda \text{ is any index set}$$

$$\therefore \quad X = f^{-1}(Y_1) = f^{-1}\left(\bigcup_{\alpha \in \Lambda} V_\alpha\right) = \bigcup_{\alpha \in \Lambda} f^{-1}(V_\alpha)$$

Since V_α is open in Y_1, and f is continuous.

$\therefore$ $f^{-1}(V_\alpha)$ is open in X. Hence $\{f^{-1}(V_\alpha)\}$ is an open cover of X. But X is compact, therefore there exists a finite subcover, say

$\{f^{-1}(V_{\alpha_1}), f^{-1}(V_{\alpha_2}), \ldots, f^{-1}(V_{\alpha_n})\}$ of the open cover $\{f^{-1}(V_\alpha)\}$ of X such that

$$X = \bigcup_{i=1}^{n} f^{-1}(V_{\alpha i})$$

Let $y \in Y_1 = f(X)$. Then there exists an $x \in X$ such that

$$y = f(x).$$

Since $x \in X$ and $X = \bigcup_{i=1}^{n} f^{-1}(V_{\alpha_i})$, $\therefore x \in \bigcup_{i=1}^{n} f^{-1}(V_{\alpha_i})$, and hence $x \in f^{-1}(V_{\alpha_i})$, for some i

or $$f(x) \in V_{\alpha_i}, \text{ for some } i$$

i.e., $$y \in V_{\alpha_i}, \text{ for some } i$$

$$\therefore \quad Y_1 = \bigcup_{i=1}^{n} V_{\alpha_i}$$

$\therefore$ $\{V_{\alpha_1}, V_{\alpha_2}, \ldots, V_{\alpha_n}\}$ is a finite subcover of the open cover $\{V_\alpha\}$ of Y_1. Therefore, Y_1 is compact.

If f is onto then $Y_1 = Y$, and so Y is compact.

Example 40. Let A be a non-empty compact subset of a metric space (X, d) and let F be a closed subset of X such that $A \cap F = \phi$, then $d(A, F) > 0$.

- If possible $d(A, F) = 0$. Since the function $x \to d(x, F)$ is continuous on A, and A being compact implies $d(x, F)$ assumes a minimum value for some $x \in A$, say x_0. And so

$$d(x_0, F) = d(A, F) = 0$$

This implies $x_0 \in \overline{F} = F$

Hence $x_0 \in A \cap F$, *i.e.*, $A \cap F \neq \phi$, which is a contradiction.

5.1 Compactness and Finite-Intersection Property

Definition. A family of subset of a non-empty set X is said to have the *finite-intersection property* (*FIP*) if every finite subfamily has non-empty intersection.

Ex. The family $\{[-n, n] : n \in \mathbf{N}\}$ of closed (intervals) subsets of $\mathbf{R}$ has the FIP.

Theorem 23. *The metric space (X, d) is compact if and only if every family of closed sets in (X, d) with the FIP has non-empty intersection.*

Let (X, d) be compact, and let $\{F_\alpha\}$ be any family of closed sets in (X, d) with FIP. If possible, let $\bigcap_{\alpha \in \Lambda} F_\alpha$ is empty, then on taking complements in X, we get

$$\bigcup_{\alpha \in \Lambda} F_\alpha^c = X$$

Thus the collection $\{F_\alpha^c\}$ of open sets, being complements of closed sets F_α in (X, d), is an open cover of the compact metric space X ; which has a finite subcover say $F_{\alpha_1}^c, F_{\alpha_2}^c, \ldots, F_{\alpha_n}^c$

i.e.,
$$\bigcup_{i=1}^{n} F_{\alpha_i}^c = X$$

Taking complements

$$\bigcap_{i=1}^{n} F_{\alpha_i} = \phi$$

which is a contradiction to the fact that $\{F_\alpha\}$ has the FIP. Hence $\bigcap_{\alpha \in \Lambda} F_\alpha$ is non-empty.

Conversely, suppose every family of closed sets in (X, d) with the FIP has non-empty intersection. In order to show that (X, d) is compact, let $\{G_\alpha\}$ be an open cover of X. Then $\bigcup_{\alpha \in \Lambda} G_\alpha = X$

Taking complements

$$\bigcap_{\alpha \in \Lambda} G_\alpha^c = \phi$$

Therefore $\{G_\alpha^c\}$ is a family of closed sets in X, whose intersection is empty. Therefore, by hypothesis this family does not have the FIP and so there exists a finite subfamily, say $G_{\alpha_1}^c, G_{\alpha_2}^c \ldots G_{\alpha_n}^c$ such that

$$\bigcap_{i=1}^{n} G_{\alpha_i}^c = \phi$$

i.e.,
$$\left(\bigcup_{i=1}^{n} G_{\alpha_i}\right)^c = \phi, \text{ or } \bigcup_{i=1}^{n} G_{\alpha_i} = X$$

Hence $\{G_{\alpha_i} : i = 1, 2, \ldots, n\}$ is a finite subcover of $\{G_\alpha\}$, and so (X, d) is compact.

Definition. A metric space (X, d) is said to have *Bolzano-Weierstrass Property* if every infinite subset of X has a limit point.

The space **R** with the usual metric does not have Bolzano-Weierstrass property for the set $\{1, 2, 3, \ldots,\}$ is an inifinite set in **R** with no limit points.

Definition. A metric space (X, d) is *sequentially compact* if every sequence $\{x_n\}$ in X has a convergent subsequence.

Theorem 24. *A metric space (X, d) is sequentially compact if and only if it has the Bolzano-Weierstrass property.*

Let (X, d) be sequentially compact metric space. Let A be an infinite subset of X. We shall show that A has a limit point. Since A is infinite, we can always extract a sequence say $\{a_n\}$ of distinct points from A. Since X is sequentially compact therefore $\{a_n\}$ contains a convergent subsequence $\{a_{n_k}\}$.

Let $\lim\limits_{k\to\infty} a_{nk} = a$. Consider any neighbourhood $S_\varepsilon(a)$ of 'a'. Then $\varepsilon > 0$ implies there exists a positive integer m such that

$$d(a_{n_k}, a) < \varepsilon, \quad \forall\ k \geq m$$

$$\Rightarrow \qquad a_{n_k} \in S_\varepsilon(a), \quad \forall\ k \geq m$$

This implies a is a limit point of A.

Conversely, suppose (X, d) has the Bolzano-Weierstrass property. Let $\{x_n\}$ be an arbitrary sequence in (X, d) and $S = \{x_n : n \in \mathbf{N}\}$ be its range. There are two possibilities—either S is finite or infinite.

(*i*) If S is finite, then there must exist at least one number $x \in S$ such that $x_n = x$ for infinitely many values of n and so the sequence $\{x_n\}$ has a constant subsequence and hence convergent.

(*ii*) When S is infinite, then by hypothesis S has a limit point say x_0. Therefore for each $\varepsilon > 0$, the set $S \cap S_\varepsilon(x_0)$ is infinite. Choose

$x_{n_1} \in S \cap S_1(x_0),\ x_{n_2} \in S \cap S_{1/2}(x_0),\ x_{n_3} \in S \cap S_{1/3}(x_0), \ldots$ and so on.

Having chosen $x_{n_1}, x_{n_2}, x_{n_3}, \ldots, x_{n_k}$ choose $x_{n_{k+1}} \in S \cap S_{1/k+1}(x_0)$ with $n_{k+1} > n_k$. The subsequence $\{x_{n_k}\}$ of $\{x_n\}$ converges to x_0 since $x_{n_k} \in S \cap S_{1/k}(x_0)$ which implies $d(x_{n_k}, x_0)\ 1/k$ and hence (X, d) is sequentially compact.

Theorem 25. *Every compact metric space (X, d) is sequentially compact.*

Suppose A is an infinite subset of S which has no limit point in X. Then for each $x \in A$, there is an $\varepsilon_x > 0$ such that $S_{\varepsilon_x}(x) \cap A = \{x\}$. Otherwise x would be a limit point of A. Clearly the family of sets, $\{S_{\varepsilon_x}(x) : x \in \cap A\} \cup \{X - A\}$ is an open cover of X which admits no finite subcover, this contradicts the compactness of X. Hence A must have a limit point in X. Therefore by the above theorem (X, d) is sequentially compact.

5.2 Relative Compactness, ε-Nets and Totally Bounded Sets

Definition. A subset A of a metric space (X, d) is said to be *relatively compact* if $\overline{A}$ is compact.

We have seen that compact sets are always closed, so we can say that compact sets are relatively compact.

ε-Net

Let A be a subset of the metric space (X, d). Let ε be a positive real number. Then by an ε*-net* of A we mean a non-empty subset B of A such that for any $a \in A$, there exists a point $x \in B$ with $d(a, x) < \varepsilon$.

In other words each point in A comes within ε-distance of one of the points in the set B.

For example suppose $A = \mathbf{R}^2$. Then the set $B = \{(m, n): m, n = 0, \pm 1, \pm 2, \pm 3...\}$ constitute an ε*-net* for $\mathbf{R}^2$ provided $\varepsilon > \dfrac{\sqrt{2}}{2}$.

It is easy to see that a set is bounded if and only if it has an ε*-net*.

Definition. A non-empty subset 'A' of a metric space (X, d) is said to be totally bounded if for any $\varepsilon > 0$ there exists a finite ε*-net* for A, *i.e.*, if for every $\varepsilon > 0$, there is a finite number of open spheres of radius ε whose union is A.

i.e.,
$$A = \bigcup_{x \in B} S_\varepsilon(x),$$

where $\boldsymbol{B}$ is a finite ε*-net* for $\boldsymbol{A}$. Clearly total boundedness implies boundedness. Since a totally bounded set is the union of a finite number of bounded sets (open spheres). But the converse is not always true. In the case of Euclidean spaces the converse also holds. In general this is not so as can be seen by the following examples.

Example 41. Infinite discrete space X is bounded but not totally bounded, for it has no finite $\frac{1}{2}$-net, since

$$S_{1/2}(x) = \{x\},\ x \in X \text{ and } X \text{ is infinite.}$$

Example 42. Consider the space l_2 consisting of sequences $\{x_n\}$ of complex numbers such that

$$\sum_{n=1}^{\infty} |x_n|^2 < \infty,$$

and the metric is defined by

$$d(x, y) = \left(\sum_{n=1}^{\infty} |x_n - y_n|^2\right)^{1/2}, \text{ where } x = \{x_n\},\ y = \{y_n\} \in l_2$$

- Let A be a subset of l_2 consisting of sequences

$$e_1 = (1, 0, 0, \ldots),\quad e_2 = (0, 1, 0, \ldots, 0),\quad e_3 = (0, 0, 1, 0, 0, \ldots, 0)$$

since $d(e_i, e_j) = \sqrt{2}, \forall\, i \neq j$, therefore A is bounded, we shall show that A is not totally bounded. Observe that A has no finite $\dfrac{1}{\sqrt{2}}$-net, for if it has, then there exists a finite set B of X such that

$$d(e_i, x) < \frac{1}{\sqrt{2}}, \text{ and } d(e_j, y) < \frac{1}{\sqrt{2}}, \text{ for } i \neq j, \text{ and } x, y \text{ in } B$$

Clearly $x \neq y$, for $x = y$ implies by triangle inequality

$$\sqrt{2} = d(e_1, e_j) \leq d(e_i, x) + d(x, e_j) < \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}} = \sqrt{2}$$

So for each e_i in A there is an x in B with the above property. Thus there corresponds an infinite set B, which is a contradiction to the fact that B is finite.

Ex. Every subset of a totally bounded set is totally bounded.

Note: Every compact metric space (X, d) is totally bounded, since for each $\varepsilon > 0$, the open cover $\{S_\varepsilon(x) : x \in X\}$ of X has a finite subcover.

Note that infinite discrete space and the set A in the above examples are not compact.

Recall that a metric space is separable if it has a countable dense subset. We have the following :

Theorem 26. *Every totally bounded metric space (X, d) is separable.*

Since X is totally bounded, therefore for each positive integer n it has a finite $1/n-$net, say A_n. Then A_n is a finite set, and

$$X = \bigcup_{a \in A_n} S_{1/n}(a)$$

Let $A = \bigcup_{n=1}^{\infty} A_n$. A is a countable subset of X, being a countable union of finite sets A_n. We shall now prove that A is dense in X, *i.e.*, $\overline{A} = X$. For this, let x be any element of X. Let $S_\varepsilon(x)$ be any open sphere centered at x. Choose a positive integer n such that $\frac{1}{n} < \varepsilon$. Since A_n is $\frac{1}{n}$-net, therefore $x \in X = \bigcup_{a \in A_n} S_{1/n}(a)$ implies $x \in S_{1/n}(a)$ for some $a \in A_n$. This implies

$$d(x, a) < \frac{1}{n} < \varepsilon$$

i.e., $$d(x, a) < \varepsilon, \text{ and so } a \in S_\varepsilon(x)$$

Therefore $S_\varepsilon(x) \cap A_n \neq \phi$, and hence $S_\varepsilon(x) \cap A \neq \phi$

$\therefore$ $x \in \overline{A}$. This shows that A is a dense subset of X.

Corollary. *Every compact metric space is separable.*

Note: However, the converse may not be true. For example (R, d) is a separable but not compact.

The next theorem characterizes total boundedness in terms of sequences in the space, but first we need the following lemma.

Lemma. *If A is an infinite subset of a totally bounded metric space (X, d), then for each $\varepsilon > 0$, there exists an infinite set $B \subseteq A$, such that $d(B) < \varepsilon$.*

Let $\varepsilon > 0$ be given. For $\varepsilon/3 > 0$, (X, d) being totally bounded, has a finite $\varepsilon/3$ net, say $\{x_1, x_2, \ldots, x_n\}$. Then

$$X = \bigcup_{i=1}^{n} S_{\varepsilon/3}(x_i)$$

$$\therefore \qquad A = \bigcup_{i=1}^{n} A \cap S_{\varepsilon/3}(x_i) \qquad [\because \; A \subseteq X]$$

This implies at least one of the sets $A \cap S_{\varepsilon/3}(x_i)$ is infinite ($\because$ A is infinite) call it B. Clearly $B \subseteq A$, and $d(B) < \dfrac{2\varepsilon}{3} < \varepsilon$.

Theorem 27. *A metric space (X, d) is totally bounded if and only if every sequence in X contains a Cauchy subsequence.*

Suppose (X, d) is totally bounded and let $\{x_n\}$ be any sequence in X. Let A denote the range set of the sequence. If A is finite then there is nothing to prove. Otherwise by the above Lemma $\exists$ an infinite set $B_1 \subseteq A$, such that

$$d(B_1) < 1$$

Choose a positive integer n_1 such that $x_{n_1} \in B_1$. Again by the same argument $\exists$ an infinite set $B_2 \subseteq B_1$ such that

$$d(B_2) < \frac{1}{2}$$

Choose a positive integer $n_2 > n_1$ such that $x_{n_2} \in B_2$. Continuing in this manner we obtain a subsequence $\{x_{nk}\}$ of $\{x_n\}$ such that $x_{nk} \in B_k$ with $d(B_k) < \dfrac{1}{k}$ and $B_{k+1} \subseteq B_k$, $k = 1, 2, 3. \ldots$

We shall now prove that $\{x_{nk}\}$ is Cauchy.

Let $\varepsilon > 0$ be given. Choose a positive integer k_0 such that $\dfrac{1}{k_0} < \varepsilon$

$\therefore$ For $k, m \geq k_0$ we have by our construction of B_k's; $x_{nk}, x_{nm} \in B_{k_0}$, and $d(B_{k_0}) < \dfrac{1}{k_0}$

$$\Rightarrow \qquad d(x_{n_k}, x_{n_m}) < \frac{1}{k_0} < \varepsilon$$

$\Rightarrow$ $\{x_{nk}\}$ is Cauchy. Hence every sequence in (X, d) has a Cauchy subsequence.

Conversely, suppose (X, d) is not totally bounded.

Then there exists $\varepsilon_0 > 0$, for which there is no finite ε_0-net for X.

Let $x_1 \in X$ be arbitrary. Then $S_{\varepsilon_0}(x_1) \neq X$ (for otherwise $\{x_1\}$ is an ε_0-net for X).

This implies $\exists\, x_2 \in X$ such that

$$x_2 \notin S_{\varepsilon_0}(x_1), \ \ i.e., \ \ d(x_2, x_1) \geq \varepsilon_0$$

Again we have $S_{\varepsilon_0}(x_1) \cup S_{\varepsilon_0}(x_2) \neq X$ (for otherwise $\{x_1, x_2\}$ is an ε_0-net for X), therefore there exists $x_3 \in X$ such that

$$x_3 \notin S_{\varepsilon_0}\,(x_1) \cup S_{\varepsilon_0}\,(x_2) \ \ i.e., \ d(x_3, x_1) \geq \varepsilon_0, \text{ and } d(x_3, x_2) \geq \varepsilon_0$$

Continuing in this manner we obtain a sequence $\{x_n\}$ in X such that $d(x_n, x_m) \geq \varepsilon_0$, for $n \neq m$. This implies the sequence $\{x_n\}$ is not Cauchy and so it has no Cauchy subsequence.

Note: In the metric space, totally boundedness is the property that complements completeness to guarantee sequential compactness (and hence compactness) as proved in the following theorem.

Theorem 28. *A metric space (X, d) is sequentially compact if and only if it is complete and totally bounded.*

Let (X, d) be sequentially compact, then every Cauchy sequence $\{x_n\}$ in X has a convergent subsequence and hence it must itself converge. Therefore, (X, d) is complete.

Again if $\{x_n\}$ is any sequence in X then it has a convergent subsequence and so by the above theorem (X, d) is totally bounded.

Conversely, suppose (X, d) is complete and totally bounded. Let $\{x_n\}$ be any sequence in X then totally boundedness of X implies $\{x_n\}$ has a Cauchy subsequence say $\{x_{n_k}\}$. But since (X, d) is complete, therefore the sequence $\{x_{n_k}\}$ must converge and hence (X, d) is sequentially compact.

Example 43. The subspace $X =]0, 1[$ of the real line is totally bounded but certainly not sequentially compact, for consider the sequence $\{1/n\}$, which has no convergent subsequence.

Note than X is not complete, since it is not closed.

Example 44. A subset A of a metric space (X, d) is totally bounded if and only if $\overline{A}$ is totally bounded.

- Let A be totally bounded. To show that $\overline{A}$ is totally bounded, it is enough to show that every sequence in $\overline{A}$ contains a Cauchy subsequence. Let $\{x_n\}$ be any sequence in $\overline{A}$. Let $\varepsilon > 0$ be given. Then $x_n \in \overline{A}$ implies

$$S_{\varepsilon/3}(x_n) \cap A \neq \phi$$

i.e., $\quad \exists\, a_n \in A$ such that $d(a_n, x_n) < \varepsilon/3$ ...(1)

So we obtain a sequence $\{a_n\}$ in A, and A being totally bounded implies $\{a_n\}$ contains a Cauchy subsequence say $\{a_{n_k}\}$. Therefore for $\varepsilon > 0, \exists$ a positive integer m such that

$$d(a_{n_j}, a_{n_k}) < \varepsilon/3, \quad \forall\, n_j, n_k \geq m \qquad ...(2)$$

By using triangle inequality and from (1) and (2), we have

$$\begin{aligned} d(x_{n_j}, x_{n_k}) &\leq d(x_{n_j}, a_{n_j}) + d(a_{n_j}, a_{n_k}) + d(a_{n_k}, x_{n_k}) \\ &< \varepsilon/3 + \varepsilon/3 + \varepsilon/3 = \varepsilon, \quad \forall\, n_j, n_k \geq m \end{aligned}$$

Hence, $\{x_{n_k}\}$ is a Cauchy subsequence of $\{x_n\}$.

Therefore, $\overline{A}$ is totally bounded.

The converse is obvious since A, being a subset of a totally bounded set $\overline{A}$, is itself totally bounded.

In order to show that sequential compactness implies compactness we need the notion of Lebsegue number for an open cover.

Lebesgue number for covers

Let $\{G_\alpha : \alpha \in \Lambda\}$ be an open cover of a metric space (X, d). A real number $\lambda > 0$ is said to be a *Lebesgue number* for the open cover $\{G_\alpha\}$ if for each subset A of X with $d(A) < \lambda$, there is at least one set G_α which contains A.

Note: Not every open cover of a metric space has a Lebesgue number. For example, let $X =]0, 1[$ be a subspace of the real line and $\{]1/n, 1[: n = 2, 3, 4, \ldots\}$ be an open cover of $]0, 1[$. For arbitrary $\lambda > 0$ the set $A =]0, \lambda/2[$ is such that $d(A) < \lambda$. But A is not contained in any of the members of the cover. Note that this space is not sequentially compact.

Theorem 29. ***Lebesgue covering lemma.*** *Every open cover of a sequentially compact metric space (X, d) has a Lebesgue number.*

Let $\{G_\alpha : \alpha \in \Lambda\}$ be any open cover of X. Assume that it has no Lebesgue number. Then for each natural number n there is a non-empty set $A_n \subseteq X$ with $d(A_n) < \dfrac{1}{n}$ such that

$$A_n \not\subset G_\alpha, \text{ for every } \alpha \in \Lambda$$

For each $n \in \mathbf{N}$, choose a point $a_n \in A_n$. Since X is sequentially compact, the sequence $\{a_n\}$ contains a convergent subsequence, say $\{a_{nk}\}$.

Let $\lim\limits_{k \to \infty} a_{n_k} = x$

Now since $x \in X = \bigcup\limits_{\alpha \in \Lambda} G_\alpha$ implies $x \in G_\alpha$, for some $\alpha \in \Lambda$. G_α being open, therefore there is an $\varepsilon > 0$ such that

$$S_\varepsilon(x) \subseteq G_\alpha.$$

For the above $\varepsilon > 0$, $a_{n_k} \to x$, and $d(A_{n_k}) \to 0$, as $k \to \infty$ implies there exists a positive integer k_0, such that

$$d(a_{n_{k_0}}, x) < \varepsilon/2, \text{ and } d(A_{n_{k_0}}) < \varepsilon/2 \qquad \ldots(1)$$

Let y be any element of $A_{n_{k_0}}$, then by using triangle inequality, and (1) we get

$$\begin{aligned} d(y, x) &\le d(y, a_{n_{k_0}}) + d(a_{n_{k_0}}, x) \\ &\le d(A_{n_{k_0}}) + d(a_{n_{k_0}}, x) \qquad [\because\ a_{n_{k_0}} \in A_{n_{k_0}}] \\ &< \varepsilon/2 + \varepsilon/2 = \varepsilon \end{aligned}$$

This implies that $y \in S_\varepsilon(x) \subseteq G_\alpha$. Hence $A_{n_{k_0}} \subseteq G_\alpha$, which contradicts the fact that for each natural number n, $A_n \not\subset G_\alpha$. Hence, $\{G_\alpha\}$ must have a Lebesgue number.

We are now in a position to prove the converse of the Theorem 25, which will establish the equivalence of compactness and sequential compactness in metric spaces.

Theorem 30. *Every sequentially compact metric space (X, d) is compact.*

Let $\{G_\alpha\}$ be any open cover of X. Since (X, d) is sequentially compact therefore by above lemma $\{G_\alpha\}$ has a Lebesgue number say $\lambda > 0$. Also (X, d) being sequentially compact is totally bounded and so it has a finite $\frac{\lambda}{3}$-net, say $\{x_1, x_2 \ldots x_n\}$.

Then
$$X = \bigcup_{i=1}^{n} S_{\lambda/3}(x_i)$$

Now for each i, $1 \le i \le n$ we have $d(S_{\lambda/3}(x_i)) \le \frac{2\lambda}{3} < \lambda$, and so by definition of Lebesgue number there exists at least one G_{α_i} such that

$$S_{\lambda/3}(x_i) \subseteq G_{\alpha_i},\ i = 1, 2, \ldots n.$$

This implies
$$\bigcup_{i=1}^{n} S_{\lambda/3}(x_i) \subseteq \bigcup_{i=1}^{n} G_{\alpha_i}$$

i.e.,
$$X \subseteq \bigcup_{i=1}^{n} G_{\alpha_i}$$

Hence $\{G_{\alpha_1}, G_{\alpha_2}, \ldots, G_{\alpha_n}\}$ is a finite subcover of $\{G_\alpha\}$ and so (X, d) is compact.

Corollary. *A metric space is compact if and only if it is sequentially compact.*

Theorem 31. *A metric space is compact if and only if it is complete and totally bounded.*

Follows from theorem 28 and the above Corollary.

Theorem 32. *A closed subspace of a complete metric space is compact if and only if it is totally bounded.*

Since a closed subspace of a complete metric space is itself complete, result follows from the above theorem.

We have seen that compactness is another name of Heine-Borel Property. Our results so far establish the following equivalence in a metric space.

(1) Bolzano-Weierstrass Property
(2) Compactness
(3) Sequential compactness
(4) Completeness and totally boundedness.

As a consequence of the Lebesgue covering lemma and the above corollary, we have the following useful result.

Theorem 33. *Let f be a continuous function from a compact metric space (X, d_1) into a metric space (Y, d_2). Then f is uniformly continuous.*

Let $\varepsilon > 0$ be given. For each x in X, $f^{-1}(S_{\varepsilon/2}(f(x)))$ is an open subset of X containing x, being an inverse image of an open sphere $S_{\varepsilon/2}(f(x))$ in Y under the continuous function $f : X \to Y$.

Therefore, the collection $\{f^{-1}(S_{\varepsilon/2}(f(x)))\}$ is an open cover of X. Since X is compact, therefore by the Lebesgue covering lemma and above corollary, this open cover has a Lebesgue number, say, $\delta > 0$. Let x, y be any two elements of X with $d_1(x, y) < \delta$, then the set $\{x, y\}$ is a set in X with diameter less than δ and so by the definition of Lebesgue number $x, y \in f^{-1}(S_{\varepsilon/2}(x'))$ for some $x' \in X$

i.e., $$f(x), f(y) \in S_{\varepsilon/2}(f(x')),$$

$$\Rightarrow \quad d_2(f(x), f(x')) < \varepsilon/2, \text{ and } d_2(f(y), f(x')) < \varepsilon/2$$

By triangle inequality,

$$d_2(f(x), f(y)) \le d_2(f(x), f(x')) + d_2(f(x'), f(y)) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

Hence f is uniformly continuous.

Aliter.

Let $\varepsilon > 0$ be given. f being continuous on X, implies for each $x \in X$; there exists $\delta_x > 0$ (depending on ε and x) such that for $x' \in X$

$$d_1(x', x) < \delta_x \Rightarrow d_2(f(x'), f(x)) < \varepsilon/2 \qquad ...(1)$$

The collection $\{S_{\delta_x/2}(x) : x \in X\}$ of open spheres forms an open cover of the compact space X, therefore there exists a finite subcover $\{S_{\delta_i/2}(x_i) : x_i \in X, 1 \le i \le n\}$ such that

$$X = \bigcup_{i=1}^{n} S_{\delta_i/2}(x_i) \qquad ...(2)$$

Choose $\delta = \min\left\{\frac{\delta_1}{2}, \frac{\delta_2}{2}, \ldots, \frac{\delta_n}{2}\right\}$, then $\delta > 0$; since each $\delta_i > 0$.

Let x, y be any two elements of X, with $d_1(x, y) < \delta$. Then by (2),

$$x \in S_{\delta_i/2}(x_i) \text{ for some } i.$$

Therefore, using (1) we have

$$d_1(x, x_i) < \frac{\delta_i}{2} < \delta_i \Rightarrow d_2(f(x), f(x_i)) < \varepsilon/2 \qquad ...(3)$$

Also, by triangle inequality, we have

$$d_1(x_i, y) \le d_1(x_i, x) + d_1(x, y) < \frac{\delta_i}{2} + \delta < \frac{\delta_i}{2} + \frac{\delta_i}{2} = \delta_i$$

(by the choice of δ)

and so, again from (1)

$$d_2(f(x_i), f(y)) < \varepsilon/2 \qquad ...(4)$$

From equations (3) and (4), using triangle inequality, we obtain

$$d_2(f(x), f(y)) < d_2(f(x), f(x_i)) + d_2(f(x_i), f(y))$$
$$< \varepsilon/2 + \varepsilon/2 = \varepsilon.$$

Hence f is uniformly continuous.

Definition. A collection **F** of functions from the metric space (X, d_1) to the metric space (Y, d_2) is called *equicontinuous*, if for each $\varepsilon > 0$, there exists a $\delta > 0$, such that

$$d_2(f(x), f(y)) < \varepsilon, \text{ whenever } d_1(x, y) < \delta, \text{ for all } x, y \in X, \text{ and all } f \in \mathsf{F}.$$

According to this definition, the functions belonging to the equicontinuous collection are uniformly continuous.

EXERCISE

1. If A and B are two compact subsets of a metric space (X, d). Prove that $A \cup B$ and $A \cap B$ are compact.
2. If A and B are non-empty subsets of a metric space (X, d) and B is compact. Prove that $d(A, B) = 0$ if and only if $\overline{A} \cap B$ is non-empty.
3. Let (X, d) be any metric space and let A and B are subsets of X. Prove that if A is closed and B is compact and $d(A, B) = 0$, then $A \cap B \neq \phi$.
4. If A is a compact set of diameter $d(A)$. Prove that there exists a pair of points x and y of A such that $d(x, y) = d(A)$.
5. If A and B are disjoint compact sets in a metric space (X, d), then prove that $d(A, B) > 0$. Show also that there exists disjoint open sets G_1 and G_2 such that $A \subseteq G_1$, $B \subseteq G_2$.
6. Prove that a metric space (X, d) is compact if and only if every family of closed sets with an empty intersection has a finite subfamily with empty intersection.
7. If $\{F_\alpha : \alpha \in \Lambda\}$ is an infinite family of closed sets with the finite intersection property, and one of the sets of the family is compact. Prove that $\bigcap_{\alpha \in \Lambda} F_\alpha$ is not empty.
8. Let $X = \{1, \frac{1}{2}, \frac{1}{3}, ...\}$, and let d be the Euclidean metric. Show that the set $\{1, \frac{1}{3}, \frac{1}{5}, ...\}$, is closed and bounded set in (X, d) but not compact. Explain why this does not contradict Heine-Borel Theorem.
9. Show that a subspace of $\mathbf{R}^n$ is bounded if and only if it is totally bounded.
10. If A is a subspace of a complete metric space, show that $\overline{A}$ is compact if and only if A is totally bounded.
11. Prove *Ascoli's theorem*: A closed subspace **F** of $C[0, 1]$ is compact if and only if **F** is equicontinuous and uniformly bounded.
12. Show that a closed subspace of a complete metric space is compact if and only if it is totally bounded.
13. Prove that a metric space (X, d) is bounded if and only if it has an ε-net, for some $\varepsilon > 0$.
14. Prove that boundedness and total boundedness are equivalent in Euclidean spaces.
15. Let (X, d) be a compact metric space and let (Y, d') be any metric space. Prove that if $f: X \to Y$ is one to one continuous and onto, then f is a homeomorphism.
16. Prove that the set of all functions which are continuous and nowhere differentiable on $[0, 1]$ is a set of the second category in the space $C[0, 1]$.
17. Prove that from any infinite open cover of a separable metric space one can extract a countable open cover.
18. Prove that a separable metric space is compact if from every countable open cover, one can extract a finite open cover.
19. Let $A = \mathbf{N} \times \mathbf{N}$, and set

$$F_{(m,n)} = \{(x, y) : x, y \in \mathbf{R}, \text{ and } |x| > m, |y| > n\}.$$

Show that $\{F_{(m,n)}\}$ has the finite intersection property, and further show that $\cap \{F_{(m,n)}\} = \phi$.

20. Show that a metric space X is compact if and only if every real-valued continuous function on X is bounded.

21. Let $X = \{x: 0 < d(0, x) \le 1,$ and $x \in \mathbf{R}^2\}$, where $0 = (0, 0)$, and d is usual metric. Show that X is closed and bounded, but not compact. Also show that X is not totally bounded.

6. CONNECTEDNESS

So far, we have discussed the three important C's in metric spaces viz. the continuity, completeness, and compactness. The fourth important C which plays the vital role as regards the separation or connection between the subsets of a metric space is *connectedness*. The word connected means not separated so let us first define what we mean by separated sets.

Separated Sets

Two sets A and B in a metric space (X, d) are said to be *separated* if neither has a point in common with the closure of the other.

i.e., $$A \cap \overline{B} = \phi, \text{ and } \overline{A} \cap B = \phi$$

Note that if A and B are separated then they are disjoint since

$$A \cap B \subseteq A \cap \overline{B} = \phi,$$

but two disjoint sets are not necessarily separated. For example if,

$$A = \{x : -\infty < x < 0\},\ B = \{x : 0 \le x < \infty\},$$

then A and B are disjoint but not separated. Clearly subsets of two separated sets are themselves separated. Two closed sets (open sets) are separated if and only if they are disjoint.

If the union of two separated sets is a closed set (open set) then the sets are themselves closed (open). For if A and B are two separated sets such that the $A \cup B$ is closed, then

$$\overline{A} = \overline{A} \cap (\overline{A} \cup \overline{B}) = \overline{A} \cap \overline{(A \cup B)} = \overline{A} \cap (A \cup B) = A \cup \phi = A$$

Definition. A subset A of a metric space (X, d) is said to be *connected* if it cannot be expressed as the union of two non-empty separated sets. If A is not connected, then it is said to be *disconnected*.

Any discrete space with more than one point is disconnected. Equivalent definitions for connectedness are contained in the following theorem.

Theorem 34. *Let Y be a subset of a metric space (X, d), then the following are equivalent :*

(*i*) *Y is connected.*

(*ii*) *Y cannot be expressed as disjoint union of two non-empty closed sets in Y.*

(*iii*) *ϕ and Y are the only sets which are both open and closed in Y.*

(*i*) $\Rightarrow$ (*ii*). If possible, let $Y = A \cup B$, where A and B are closed in Y and $A \ne \phi$, $B \ne \phi$, $A \cap B = \phi$. We claim that A and B are separated.

i.e., $$A \cap \overline{B} = \phi, \text{ and } \overline{A} \cap B = \phi.$$

Clearly $Y \cap \overline{A} = A$, and $Y \cap \overline{B} = B$ ($\because$ A and B are closed in Y).

$$\therefore \quad A \cap \overline{B} = (\overline{A} \cap Y) \cap \overline{B} = \overline{A} \cap (Y \cap \overline{B}) = \overline{A} \cap B.$$

If possible, let $A \cap \overline{B} \ne \phi$, therefore there exists

$$y \in A \cap \overline{B}$$

$\Rightarrow$ $$y \in A, \text{ and } y \in \overline{B}$$

But since $A = Y \cap \overline{A}$

$\therefore$ $$y \in Y$$

Now $y \in \overline{B}, y \in Y$ implies every neighbourhood of y in Y intersects B and so $y \in B$ ($\because$ B is closed in Y).

Thus $y \in A \cap B$, implies $A \cap B \neq \phi$, which is a contradiction.

Hence $A \cap \overline{B} = \phi$.

Thus Y is the union of two non-empty separated sets. This implies that Y is disconnected, which is a contradiction to the given hypothesis. Hence (*ii*) is true.

(*ii*) $\Rightarrow$ (*i*). If possible, let Y be disconnected, then $Y = A \cup B$, where A and B are two non-empty subsets of Y such that

$$A \cap \overline{B} = \overline{A} \cap B = \phi$$

Clearly $A \cap B = \phi$ $\qquad (\because \ A \cap B \subseteq A \cap \overline{B} = \phi)$

Now $$Y \cap \overline{A} = (A \cup B) \cap \overline{A} = (A \cap \overline{A}) \cup (B \cap \overline{A}) = A \cup \phi = A$$

$$[\because \ B \cap \overline{A} = \phi]$$

Therefore A, being the intersection of Y and the closed set $\overline{A}$, is closed in Y.

Similarly, $Y \cap \overline{B} = B$ implies B is closed in Y.

This gives that Y is a disjoint union of two non-empty closed sets A and B in Y, which is a contradiction to (*ii*). Hence Y is connected.

(*ii*) $\Rightarrow$ (*iii*). If possible, let there exist a non-empty proper subset A of Y which is both open and closed in Y. Then its complement $B = Y - A$ in Y is both closed and open in Y and $Y = A \cup B$, $A \neq \phi$, $B \neq \phi$, $A \cap B = \phi$; which contradicts (*ii*). Hence (*iii*) must be true.

(*iii*) $\Rightarrow$ (*ii*). Obvious.

Note: The above theorem still holds if the closed sets in (*ii*) are replaced by open sets, that is, Y is connected if and only if it cannot be expressed as the disjoint union of two non-empty open sets in Y.

Theorem 35. *A subset Y of a metric space X is disconnected if and only if $Y \subseteq G_1 \cup G_2$, where G_1 and G_2 are open sets in X such that $Y \cap G_1 \neq \phi$, $Y \cap G_2 \neq \phi$, $G_1 \cap G_2 \cap Y = \phi$.*

Let Y be disconnected, then there exists a non-empty proper subset A of Y which is both open and closed in Y. This implies that its complement $B = Y - A$ is a non-empty proper subset of Y which is both closed and open in Y.

Since A and B are open in Y, therefore, there exists two open sets G_1 and G_2 in Y such that

$$A = G_1 \cap Y, \ B = G_2 \cap Y$$

$\therefore$ $$Y = A \cup B \subseteq G_1 \cup G_2$$

Also $(G_1 \cap G_2) \cap Y = (G_1 \cap Y) \cap (G_2 \cap Y) = A \cap B = \phi$

Clearly

$$G_1 \cap Y \neq \phi, G_2 \cap Y \neq \phi. \qquad [\because A \neq \phi, B \neq \phi]$$

Conversely, if there exists two open subsets G_1 and G_2 of X such that

$$Y \subseteq G_1 \cup G_2$$

and $(G_1 \cap G_2) \cap Y = \phi, G_1 \cap Y \neq \phi, G_2 \cap Y \neq \phi$

Then $Y = Y \cap (G_1 \cap G_2) = (Y \cap G_1) \cup (Y \cap G_2)$

Let $A = G_1 \cap Y,$ and $B = G_2 \cap Y,$

then A and B are open in Y.

So that $Y = A \cup B.$

Moreover $A \cap B = (G_1 \cap Y) \cap (G_2 \cap Y) = \phi$

Thus $A = Y - B$ is both open and closed in Y. Clearly A is a non-empty proper subset of Y. ($\because A \neq \phi, B \neq \phi$)

Hence Y is disconnected.

Note: Corresponding result also holds using closed sets, *i.e.*, Y is disconnected if and only if $Y \subseteq F_1 \cup F_2$, where F_1 and F_2 are closed sets in X such that

$$Y \cap F_1 \neq \phi, Y \cap F_2 \neq \phi \text{ and } F_1 \cap F_2 \cap Y = \phi$$

Remark: By the above theorem $Y \subseteq X$ is disconnected if there exists two open subsets G_1 and G_2 of X such that $Y = (G_1 \cap Y) \cup (G_2 \cap Y), G_1 \cap G_2 \subseteq Y^c, G_1 \cap Y \neq \phi$

and $G_2 \cap Y \neq \phi.$

Then we say that $\{G_1 \cap Y, G_2 \cap Y\}$ is a disconnection of Y. Note that this disconnection is not unique. It follows that Y is disconnected if and only if it has a disconnection. Similar remark hoids for closed sets.

Theorem 36. *Let A be a connected subset of a metric space X, and let B be a subset of X such that $A \subseteq B \subseteq \overline{A}$, then B is also connected.*

If possible, B be disconnected, then there exist two open subsets G and H of X such that

$$B \subseteq G \cup H, G \cap H \cap B = \phi, G \cap B = \phi, \text{ and } H \cap B \neq \phi$$

Now $A \subseteq B$ implies that $A \subseteq G \cup H$, and $G \cap H \subseteq B^c \subseteq A^c$

i.e., $G \cap H \cap A = \phi$

Also $G \cap A \neq \phi$, for it not, then $A \subseteq G^c$.

This implies that $\overline{A} \subseteq G^c$ $\qquad$ [$\because G^c$ is closed being complement of open set G]

i.e., $B \subseteq G^c$

or $B \cap G = \phi$, which is not possible.

Similarly, $H \cap A \neq \phi$

It follows that $\{G \cap A, H \cap A\}$ is a disconnection of A which is a contradiction, since A is connected.

Hence B must be connected. In particular, $\overline{A}$ is also connected.

Note: Like compactness, connectedness is preserved under continuous functions.

Theorem 37. *Continuous image of a connected set is connected.*

Let $f : X \to Y$ be a continuous function from a metric space X to a metric space Y. Let A be a connected subset of X. If $A = \phi$, then there is nothing to prove. Let $A \neq \phi$. We have to show that $f(A)$ is connected. If possible, let $f(A)$ be disconnected. Then Y contains open subsets G_1, G_2 which intersect $f(A)$ and are such that

$$f(A) \subseteq G_1 \cup G_2 \text{ and } G_1 \cap G_2 \cap f(A) = \phi$$

This implies

$$A \subseteq f^{-1}(G_1 \cup G_2) = f^{-1}(G_1) \cup f^{-1}(G_2), \text{ and}$$

$$f^{-1}(G_1 \cap G_2 \cap f(A)) = f^{-1}(\phi) = \phi$$

or

$$A \subseteq f^{-1}(G_1) \cup f^{-1}(G_2), \text{ and } f^{-1}(G_1) \cap f^{-1}(G_2) \cap A = \phi$$

$$(\because\ f^{-1}(f(A)) = A)$$

But $f^{-1}(G_1)$ and $f^{-1}(G_2)$ are both open in X, being inverse images of open sets G_1 and G_2 under the continuous function f.

Also $$f^{-1}(G_1) \cap A = f^{-1}(G_1) \cap f^{-1}(f(A)) = f^{-1}(G_1 \cap f(A)) \neq \phi$$

$$[\because\ G \cap f(A) \neq \phi]$$

Similarly $$f^{-1}(G_2) \cap A \neq \phi$$

This implies that A is disconnected, which is a contradiction.

Hence $f(A)$ must be connected.

Theorem 38. *The union of two connected sets, having non-empty intersection, is connected.*

Let A and B be any two connected subsets of a metric space X, such that

$$A \cap B \neq \phi.$$

Let $Y = A \cup B$. To show Y is connected, let D be any non-empty subset of Y which is both open and closed in Y. Then $D \subseteq Y = A \cup B$ implies D must intersect A or B or both. Suppose it intersects A, *i.e.*, $D \cap A \neq \phi$. Then $D \cap A$ is a non-empty subset of A which is both open and closed in A. Therefore $D \cap A = A$, *i.e.*, $A \subseteq D$, since A is connected. Now $A \cap B \subseteq D \cap B$, and $A \cap B \neq \phi$, so that $D \cap B$ is a non-empty subset of B which is both open and closed in B and so $D \cap B = B$, *i.e.*, $B \subseteq D$, since B is connected.

Thus, $$Y = A \cup B \subseteq D \cup D = D$$

Hence, $Y = D$. It follows that Y is connected.

Theorem 39. *An arbitrary union of connected sets, with non-empty intersection, is connected.*

Let $\{A_\alpha\}$ be a family of connected sets having non-empty intersection *i.e.*, $\underset{\alpha}{\cap} A_\alpha \neq \phi$. Taking $Y = \underset{\alpha}{\cup} A_\alpha$, the proof follows by the same argument as given in the above theorem.

Ex. 1. If $\{A_n\}$ is a sequence of connected subsets of a metric space X, each of which intersects its successor, *i.e.*, $A_n \cap A_{n+1} \neq \phi$, $\forall\ n \in \mathrm{N}$, then show that $\overset{\infty}{\underset{n=1}{\cup}} A_n$ is connected.

[*Hint*: Taking $B_n = A_1 \cup A_2 \cup A_3 \ldots \cup A_n$, we have

$$\overset{\infty}{\underset{n=1}{\cap}} B_n \supseteq A_1 \text{ and } \overset{\infty}{\underset{n=1}{\cup}} B_n = \overset{\infty}{\underset{n=1}{\cup}} A_n.]$$

Ex. 2. Show that the union of any non-empty—family $\{A_\alpha\}$ of connected subsets of a metric space X, each pair of which intersects, is connected.

[*Hint*. Fix α_0, taking $B_\alpha = A_{\alpha_0} \cup A_\alpha$ for each α, we have

$$\underset{\alpha}{\cap} B_\alpha \supseteq A_{\alpha_0} \text{ and } \underset{\alpha}{\cup} B_\alpha = \underset{\alpha}{\cup} A_\alpha.]$$

Example 45. Discuss the connectedness of the following subsets of the Euclidean space $\mathbf{R}^2$.

(*i*) $D = \{(x, y) : x \neq 0, \text{ and } y = \sin 1/x\}$

(*ii*) $E = \{(x, y) : x = 0, \text{ and } -1 \leq y \leq 1\} \cup D$

■ (*i*) $D \subseteq \mathbf{R}^2$, where $D = \{(x, y) : x \neq 0, \text{ and } y = \sin 1/x\}$

Let $A = \{(x, y) : x > 0, \text{ and } y = \sin 1/x\}$, and

$$B = \{(x, y) : x < 0, \text{ and } y = \sin 1/x\}$$

Then $$D = A \cup B, \text{ and } A \cap B = \phi$$

Since $$A = D \cap \{x, y) : x > 0\}, \text{ and}$$
$$B = D \cap \{x, y) : x < 0\}$$

The sets $\{(x, y) : x > 0\}$ and $\{(x, y) : x < 0\}$ are pairwise disjoint open subsets in $\mathbf{R}^2$. Obviously A and B are open in D and they are also non-empty, *i.e.*,

$$A \neq \phi,\ B \neq \phi$$

Therefore, $\{A, B\}$ is disconnection of D.

Hence, D is a disconnected subset of $\mathbf{R}^2$.

(*ii*) Next, we have $E \subseteq \mathbf{R}^2$, given by

$$E = D \cup \{(0, y) : -1 \leq y \leq 1\}$$

Let $F = \{(0, y) : -1 \leq y \leq 1\}$,

then $$E = A \cup B \cup F$$

From the graph of $y = \sin 1/x$, it is easy to verify that

$$\overline{A} = A \cup F, \text{ and } \overline{B} = B \cup F$$

Therefore

$$E = \overline{A} \cup \overline{B}$$

Also $$\overline{A} \cap \overline{B} = F \neq \phi$$

We now define a function $f :]0, \infty[\to \mathbf{R}^2$ by

$$f(x) = (x, \sin 1/x)$$

The function f is continuous and the set A, being the continuous image of the connected set $]0, \infty[$, is connected. So $\overline{A}$ is connected. Similarly the function $g :]-\infty, 0[\to \mathbf{R}^2$ defined by $g(x) = (x, \sin 1/x)$ is continuous. By the same argument $\overline{B}$ is connected.

Hence, E is a connected subset of $\mathbf{R}^2$.

Theorem 40. *A non-empty subset X of* $\mathbf{R}$ *(with usual metric) is connected if and only if X is an interval or a singleton.*

Let X be a non-empty connected subset of $\mathbf{R}$ containing at least two elements. If possible, let X be not an interval, then there exists $a, b, c \in \mathbf{R}$ such that

$$a < c < b,\ a, b \in X, \text{ but } c \notin X.$$

Then $G_1 =]-\infty, c[$ and $G_2 =]c, \infty[$ are two disjoint open sets in $\mathbf{R}$ which intersect X (since $a \in G_1 \cap X$ and $b \in G_2 \cap X$) such that

$$X = (X \cap G_1) \cup (X \cap G_2) \subseteq G_1 \cup G_2$$

This shows that X is disconnected, which is a contradiction. Hence X must be an interval.

Conversely, if X is a singleton set, then there is nothing to prove. If possible, let X be disconnected, then there exist two open subsets G_1 and G_2 of $\mathbf{R}$ such that

$$X \subseteq G_1 \cup G_2,\ G_1 \cap G_2 \cap X = \phi,\ G_1 \cap X \neq \phi,\ G_2 \cap X \neq \phi.$$

Let $a \in G_1 \cap X, b \in G_2 \cap X$. Assume that $a < b$. The set $[a, b] \cap G_1$ is non-empty and bounded above (by b) and so it has a supremum say ξ. Clearly $a < \xi \leq b$. Now $\xi \notin G_2$, for if it were in the open set G_2, then there exist $\varepsilon > 0$ such that

$$]\xi - \varepsilon, \xi + \varepsilon[\subseteq G_2$$

This implies $\xi - \varepsilon$ is an upper bound of $[a, b] \cap G_1$ which contradicts the choice of ξ as the least upper bound.

Similarly $\xi \notin G_1$, for if $\xi \in G_1$, G_1 being open, then there exists $\delta > 0$ such that

$$]\xi - \delta, \xi + \delta[\subseteq G_1.$$

Now we have $\xi < b$ and $]\xi, b[\cap G_2 \neq \phi$ so that ξ cannot be an upper bound of $[a, b] \cap G_1$, which contradicts the choice of ξ.

Hence $a, b \in X, a < \xi < b$ and $\xi \notin X$, it follows that X is not an interval.

Corollary. *The real line* $\mathbf{R}$ *is connected.*

Since $\mathbf{R}$ is an interval so it is connected by the above theorem.

Aliter. Let, if possible $\mathbf{R}$ be disconnected. Then there exist two non-empty disjoint closed sets A and B in $\mathbf{R}$ such that $\mathbf{R} = A \cup B$.

Let $a_1 \in A$ and $b_1 \in B$, $a_1 \neq b_1$ so either $a_1 < b_1$ or $a_1 > b_1$. Without loss of generality we may assume $a_1 < b_1$. Let $I_1 = [a_1, b_1]$. The mid-point $\frac{a_1 + b_1}{2}$ of $[a_1, b_1]$ being a point of **R** belongs to A or to B (but not to both). In case it belongs to A, take the interval $\left[\frac{a_1 + b_1}{2}, b_1\right]$ and name it $I_2 = [a_2, b_2]$; otherwise we call the interval $\left[a_1, \frac{a_1 + b_1}{2}\right]$ as $I_2 = [a_2, b_2]$, where $I_2 \subseteq I_1$ and $b_2 - a_2 = \frac{1}{2}(b_1 - a_1)$. Now bisecting I_2 as before and continuing the process indefinitely we get a sequence $\{I_n\} = \{[a_n, b_n]\}$ or closed intervals such that

$$I_1 \supseteq I_2 \supseteq I_3 \ldots \supseteq I_n \ldots$$

with the property that $a_n \in A, b_n \in B, \forall\, n \in \mathbf{N}$; and $l(I_n) = \frac{1}{2^{n-1}}(b_1 - a_1)$ which tends to zero as n tends to infinity.

Therefore by Cantor's intersection theorem,

$$\bigcap_{n=1}^{\infty} I_n = \{c\}$$

i.e.,
$$c \in I_n \quad \forall\, n \in \mathbf{N}$$

Clearly c is a limit point of both A and B, for if $\varepsilon > 0$ is any arbitrary real number then there exists a positive integer m_0 such that

$$I_n \subseteq]c - \varepsilon, c + \varepsilon[, \quad \forall\, n \geq m_0$$

This implies $]c - \varepsilon, c + \varepsilon[$ contains infinite number of points $a_{m0}, a_{m0+1} \cdots$ of A as well as infinite number of points $b_{m0}, b_{m0+1}, \cdots$ of B. So c is a limit point of both A and B. But A and B are closed subsets of **R**. Therefore, c belongs to both A and B, which is a contradiction to the fact that $A \cap B = \phi$. Hence **R** is connected.

Definition. A real-valued function is said to have an *intermediate value property* if it assumes every value between any two of its values.

Theorem 41. ***Generalized Intermediate-Value Theorem.*** *Every real-valued continuous function f defined on a connected metric space X has the intermediate-value property.*

Continuity of the function $f: X \to \boldsymbol{R}$ and connectedness of X implies $f(X)$ is a connected subset of **R**. Then from Theorem 40, $f(X)$ is either a singleton or an interval.

If $f(X)$ is singleton then there is nothing to prove. Let $f(X)$ be an interval.

Let $f(x) \neq f(y)$, $x, y \in X$ be any two values of $f(X)$. Then either $f(x) < f(y)$ or $f(x) > f(y)$ Without loss of generality we may assume that $f(x) < f(y)$.

Let A be any real number lying between $f(x)$ and $f(y)$

i.e.,
$$f(x) < A < f(y).$$

Then, $A \in f(X)$ $\qquad$ [$\because$ $f(X)$ is an interval]

i.e., $A = f(a)$, for some $x \in X$.

Hence f has the intermediate value property.

Corollary (*Intermediate-value theorem*). *If a real-valued function f is continuous on the closed interval $[a, b]$. Then f has the intermediate value property, i.e., f assumes every value between $f(a)$ and $f(b)$.*

This follows from the above theorem. Since the interval $[a, b]$ is a connected subset of R.

Theorem 42. *A metric space X is connected if and only if every real-valued continuous function f has the intermediate value property.*

The necessary part has been proved in the above theorem.

For the sufficient part we shall show that if X is not connected then there exists a real-valued continuous function which does not possess the Intermediate value property. Let X be disconnected, then there exist two disjoint non-empty open sets G and H in X such that

$$X = G \cup H$$

The function $f: X \to \boldsymbol{R}$ defined by

$$f(x) = \begin{cases} 0, & \text{if } x \in G \\ 1, & \text{if } x \in H \end{cases}$$

does not take the intermediate value $\frac{1}{2}$ at any point of X.

So f does not possess the intermediate value property. However, f is continuous. Since the range $f(X)$ of f is the discrete space $\{0, 1\}$, therefore, a complete collection of open subsets of $f(X)$ is given by ϕ, $\{0, 1\}$, $\{0\}$, $\{1\}$. By definition of f the inverse image of these sets are ϕ, X, G, H respectively, all of which are open in X.

Corollary. *A metric space X is disconnected if and only if there exists a continuous function from X onto the discrete metric space $\{0, 1\}$.*

This follows from above theorem and the fact that the discrete metric space with more than one point is always disconnected.

6.1 Components of a Metric Space

Definition. A subset A of a metric space X is said to be a *component* of X if it is the maximal connected subset of X, *i.e.,* if A is connected but not contained properly in any larger connected subset of X.

Thus if X is connected, then X is the only component of itself. Components always exist, since singleton sets are connected. Moreover, in a discrete metric space X singletons are the only connected subsets of X, since a subset A of X containing more than one element is obviously disconnected and so the components of X are singletons.

Properties:

(*i*) Components are closed sets

Let X be any metric space and A be its component. By definition, A is a maximal connected subset of X. Suppose A is not closed, then $\overline{A} \neq A$, *i.e.,* $A \subset \overline{A}$. Now $\overline{A}$, being the closure of the connected set A is itself connected which properly contains A, this contradicts the maximality of A. Hence, A must be closed.

(*ii*) Components need not be open sets

Consider the metric space of rationals $(\mathbf{Q}, d)$ with the usual metric. The components of $\mathbf{Q}$ are singletons which are not open in $\mathbf{Q}$. Let Y be any subset of $\mathbf{Q}$ containing more than one element. Choose y_1, y_2 in Y with $y_1 < y_2$, then there exists an irrational number ξ such that $y_1 < \xi < y_2$. Clearly $\{]-\infty, \xi[,]\xi, \infty[\}$ is a disconnection of Y, and so Y is disconnected.

Hence, the only connected subsets of $\mathbf{Q}$ are singletons.

(*iii*) Components of a metric space are either identical or pairwise disjoint

Let A and B be any two components of a metric space X. Then either $A \cap B = \phi$ or $A \cap B \neq \phi$. If $A \cap B = \phi$, then there is nothing to prove. If $A \cap B \neq \phi$ then $A \cup B$ is a connected subset of X. Also $A \subseteq A \cup B$ and $B \subseteq A \cup B$, the definition of components implies $A \cup B = A$, and $A \cup B = B$ and hence $A = B$.

Theorem 43. *Every connected subset of a metric space is contained in a component of X.*

Let A be any connected subset of X. Consider the collection $\{A_\alpha\}$ of all connected subsets of X containing A.

$\{A_\alpha\} \neq \phi$, since A itself is a connected subset of X containing A. Let $Y = \bigcup_\alpha A_\alpha$.

Then Y is a connected subset of X containing A, since each A_α is a connected subset of X containing A and $\bigcap_\alpha A_\alpha \neq \phi$ $(\because A \subseteq \bigcap_\alpha A_\alpha)$.

Moreover, Y is a maximal connected subset of X, for if $Y \subseteq B$, where B is a connected subset of X, then $B \in \{A_\alpha\}$ $(\because\ A \subseteq Y \Rightarrow A \subseteq B)$ and so $B \subseteq \bigcup_\alpha A_\alpha = Y$. Thus $B = Y$. This shows that Y is a maximal connected subset of X containing A; and hence Y is the required component of X containing A. Also components are pairwise disjoint, therefore Y is the only component of X containing A.

Corollary 1. *Each element of a metric space X is contained in exactly one component.*

This follows directly from the above theorem since each singleton is connected.

Corollary 2. *A non-empty connected subset of a metric space X is a component, if it is both open and closed.*

Let A be a non-empty connected subset of X, which is both open and closed. Then by the above theorem, A is contained in some component, say, B of X. To prove that A is a component of X, we shall show that $A = B$. If possible let $A \neq B$, then A is a proper subset of B which is both open and closed in B. This implies $A = \phi$ $(\because$ B being component, is connected), which is a contradiction. Hence $A = B$.

Ex. Show that the metric space if $(\mathbf{R}^n, d)$ where

$$d(x, y) = \max_{1 \le i \le n} |x_i - y_i| \quad x = (x_1, x_2, \ldots, x_n),$$

$$y = (y_1, y_2, \ldots, y_n) \in \mathbf{R}^n$$

is connected.

[*Hint*: Every non-empty proper subset of $\mathbf{R}^n$ has non-empty boundary].

20 The Lebesgue Integral

After Riemann's definition of the integral in 1854 its limitations became apparent. Many definitions of the integral for bounded as well as unbounded functions were put forward by various mathematicians. At the beginning of nineteenth century, the French mathematician Lebesgue introduced the concept of integral which widened the scope of modern analysis.

In Riemann's integration, the domain over which the integral is taken is divided into a partition, and the integral is defined as the limit of a Cauchy sum for this partition as the norm of the partition tends to zero. In Lebesgue integration, on the other hand, the domain over which the integral is taken is divided into a number of measurable sets. The integral is then defined as a limit of a certain sum taken for all these measurable sets as the number of measurable sets increases indefinitely. The basic distinction between the Lebesgue integral and the Riemann integral lies on the fact that the mode of subdivision of the domain of integration is different in both the cases. There are many ways to develop the Lebesgue integral. The development we shall give here is based on the concept of measurable sets and measurable functions. Let us first discuss measurable sets.

1. MEASURABLE SETS

Let $[a, b]$ be a closed bounded interval in **R**. By length of $[a, b]$ we mean the real number $b - a$. If A is any non-empty bounded open subset of $[a, b]$, then the length of A is defined as the sum of the lengths of all its disjoint open intervals I_k, $k = 1, 2, 3, \ldots$ *s.t.* $A \subseteq \bigcup_{k=1}^{n} I_k$, and is denoted by $l(A)$. If A_1 and A_2 are two bounded open sets and $A_1 \subseteq A_2$, then $l(A_1) \le l(A_2)$. Again if B is any closed subset of the interval $[a, b]$, the complement $C(B)$ of B relative to any open subset F of $[a, b]$ containing B is $F - B = F \cap C(B)$ and is open in $[a, b]$. Then length of B is defined as

$$l(B) = l(F) - l(F - B)$$

It is to be noted that the length of B does not depend on the choice of F containing B.

1.1 Outer and Inner Measure of a Bounded Set

Let $A \subseteq [a, b]$ be any bounded subset of real numbers. The *outer measure* of A denoted by $m^* A$, is defined as $m^* A = \inf l(F)$, where the infimum is taken over all open sets F which contain A. F being

open, can be expressed as a countable union of open intervals I_n, $n = 1, 2, \ldots$, such that $A \subseteq F \subseteq \bigcup_{n=1}^{\infty} I_n$. For each such countable collection containing A, consider the sum of the lengths of the intervals in that collection. The *outer measure* of A is then defined as

$$m^* A = \inf_{A \subseteq \bigcup_n I_n} \sum_n l(I_n)$$

It is clear that the measure of any bounded set is non-negative and finite. For an empty set ϕ, $m^* \phi = 0$, and if $A \subseteq B$, then $m^* A \le m^* B$. Moreover a set consisting of a single element has outer measure zero.

The *inner measure* of A denoted by $m_* A$ is defined as

$$m_* A = \sup l(B)$$

where the supremum is taken over the lengths of all the closed sets B contained in the set A. Thus we observe that the outer measure $m^* A$ is computed by open sets F which contain A and which come closer and closer to A.

$$\therefore \qquad m^* A \le l(F)$$

whereas the inner measure $m_* A$ is computed by closed sets B which are contained in A.

$$\therefore \qquad m_* A \ge l(B)$$

In general $m^* A \ne m_* A$.

Definition. A set $A \subseteq [a, b]$ is said to be *measurable* if $m^* A = m_* A$. In this case we define $m A$, the measure of A as

$$m A = m^* A = m_* A$$

Theorem 1. *For every set A,*

$$m_* A = m^* A$$

It suffices to prove the result if A is bounded.

Let F be any bounded open set containing A.

For any closed subset B contained in A, we have

$$B \subseteq A \subseteq F$$

$$\therefore \qquad l(B) \le l(F)$$

Taking supremum over all closed subsets B contained in A, we have

$$\sup_{B \subseteq A} l(B) \le l(F)$$

Again taking infimum over all open subsets F containing A

$$\sup_{B \subseteq A} l(B) \le \inf_{F \supseteq A} l(F)$$

i.e.,

$$m_* A \le m^* A.$$

Theorem 2. If $A \subseteq [a, b]$, *then*

$$m^* A + m_* \; C\,(A) = b - a$$

where $C\,(A) = [a, b] - A$ *is the complement of* A *relative to* $[a, b]$.

Let F be any open subset of $[a, b]$ containing A, then $C\,(F) \subseteq C\,(A)$. By definition it follows that

$$m_* \; C(A) \geq l\,(C\,(F))$$

Adding $l\,(F)$ on both sides,

$$l\,(F) + m_* \, C(A) \geq l\,(F) + l\,(C\,(F)) = (b - a)$$

Taking infimum over all open sets $F \supseteq A$

$$m^* A + m_* \, C(A) \geq (b - a) \qquad \text{...(1)}$$

Now if B is any closed set such that $B \subseteq C\,(A)$ and $C\,(B) \supseteq A$, then by definition it follows that

$$m^* A \leq l\,(C(B))$$

Adding $l\,(B)$ on both sides, we have

$$m^* A + l\,(B) \leq l\,(C\,(B)) + l\,(B) = (b - a)$$

Taking supremum over all closed sets $B \subseteq C(A)$

$$m^* A + m_* \; C(A) \leq (b - a) \qquad \text{...(2)}$$

From equations (1) and (2), we have

$$m^* A + m_* \; C(A) = (b - a)$$

Corollary. A subset $A \subseteq [a, b]$ is measurable if and only if

$$m^* A + m^* \; C(A) \leq (b - a)$$

For any set $A \subseteq [a, b]$,

$$m^* A + m_* \; C(A) = (b - a)$$

Interchanging A with $C\,(A)$, we have

$$m^* \; C(A) + m_* \,(A) = (b - a) \qquad \text{...(1)}$$

If A is measurable, then

$$m^* A = m_* \; A = m\,A \qquad \text{...(2)}$$

From equations (1) and (2), we get

$$m^* \; C(A) + m^* A = (b - a)$$

Hence, the condition is necessary.
Given

$$m^* A + m^* \; C(A) \leq (b - a) \qquad \text{...(3)}$$

Subtracting (1) from (3), we get

$$m^* A - m_* A \leq 0$$

$$\Rightarrow \qquad m^* A \leq m_* A$$

But $m_* \, A \leq m^* \, A$. Hence, $m_* \, A = m^* \, A \Rightarrow A$ is measurable.

The above result being symmetric in A and CA implies that CA is measurable whenever A is measurable.

CHAPTER 20

Ex. 1. A necessary and sufficient condition for a set A to be measurable is that for each $\in > 0$, there exists an open set $F \supseteq A$ and a closed set $B \subseteq A$ with $F \supseteq A \supseteq B$ such that $mF - mB < \varepsilon$.

Ex. 2. If A_1 and A_2 are subsets of $[a, b]$, then

$$m^* A_1 + m^* A_2 \geq m^* (A_1 \cup A_2) + m^* (A_1 \cap A_2)$$

and

$$m_* A_1 + m_* A_2 \leq m_* (A_1 \cup A_2) + m_* (A_1 \cap A_2)$$

[*Hint*: If F_1 and F_2 are any two open subsets of $[a, b]$ containing A_1 and A_2 respectively, and $F_1 \cup F_2 \supseteq A_1 \cup A_2$, $F_1 \cap F_2 \supseteq A_1 \cap A_2$, then

$$l(F_1) + l(F_2) = l(F_1 \cup F_2) + l(F_1 \cap F_2)$$

and for each $\in > 0$, $m^* A_1 + \in > l(F_1)$]

Theorem 3. *If A_1 and A_2 are measurable subsets of $[a, b]$ then both $A_1 \cup A_2$ and $A_1 \cap A_2$ are measurable and*

$$m A_1 + m A_2 = m(A_1 \cup A_2) + m(A_1 \cap A_2).$$

Now A_1, A_2 being measurable,

$$m^* A_1 = m_* A_1 \text{ and } m^* A_2 = m_* A_2$$

Thus
$$m A_1 + m A_2 = m^* A_1 + m^* A_2 \geq m^* (A_1 \cup A_2) + m^* (A_1 \cap A_2)$$
$$\geq m_* (A_1 \cup A_2) + m_* (A_1 \cap A_2)$$

and $\quad m_* (A_1 \cup A_2) + m_* (A_1 \cap A_2) \geq m A_1 + m A_2$

(using Ex. 2 and the relation of inner and outer measure).

This implies that

$$m^* (A_1 \cup A_2) + m^* (A_1 \cap A_2) = m_* (A_1 \cup A_2) + m_* (A_1 \cap A_2)$$

or
$$m^* (A_1 \cup A_2) - m_* (A_1 \cup A_2) = m_* (A_1 \cap A_2) - m^* (A_1 \cap A_2) \qquad \ldots(1)$$

But

$$m^* (A_1 \cup A_2) \geq m_* (A_1 \cup A_2)$$

$\therefore$
$$m_* (A_1 \cap A_2) - m^* (A_1 \cap A_2) \geq 0, \qquad \text{[from (1)]}$$

Again

$$m_* (A_1 \cap A_2) \leq m^* (A_1 \cap A_2)$$

$\therefore$
$$m_* (A_1 \cap A_2) = m^* (A_1 \cap A_2)$$

$\Rightarrow$ $\quad A_1 \cap A_2$ is measurable.

From (1)

$$m^* (A_1 \cup A_2) - m_* (A_1 \cup A_2) = 0$$

$\Rightarrow$ $\quad A_1 \cup A_2$ is measurable

and

$$m A_1 + m A_2 = m(A_1 \cup A_2) + m(A_1 \cap A_2)$$

Ex. 3. If A_1 and A_2 are measurable subsets of the closed interval $[a, b]$, then $A_1 - A_2$ is measurable and if $A_2 \subseteq A_1$ then $m(A_1 - A_2) = mA_1 - mA_2$.

Since $A_1 - A_2 = A_1 \cap C(A_2)$ and A_1, A_2 being measurable, $A_1 - A_2$ is also measurable. Now if $A_2 \subset A_1$ then $A_1 = A_2 \cup (A_1 - A_2)$ and $m(A_2 \cap (A_1 - A_2)) = m(\phi) = 0$. Since A_2, $A_1 - A_2$ are disjoint.

$$\therefore \quad mA_1 = m(A_2 \cup (A_1 - A_2)) + m(A_2 \cap (A_1 - A_2))$$

$$mA_1 = mA_2 + m(A_1 - A_2) \qquad \text{(by Theorem 3)}$$

or

$$mA_1 - mA_2 = m(A_1 - A_2)$$

Theorem 4. *If $A_1, A_2, \ldots, A_n$ are pairwise disjoint measurable subsets of $[a, b]$ then $\bigcup_{n=1}^{\infty} A_n$ is measurable and*

$$m\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} mA_n$$

Let $\in > 0$ be given and F_n, $n = 1, 2, 3, \ldots$ be a class of open subsets of $[a, b]$ such that $A_n \subseteq F_n$ and $l(F_n) < m^* A_n + \dfrac{\varepsilon}{2^n}$.

The set $\bigcup_{n=1}^{\infty} F_n$ is an open set containing $\bigcup_{n=1}^{\infty} A_n$.

$$\therefore \quad m^*\left(\bigcup_{n=1}^{\infty} A_n\right) \le l\left(\bigcup_{n=1}^{\infty} F_n\right) \le \sum_{n=1}^{\infty} l(F_n) < \sum_{n=1}^{\infty}\left(m^* A_n + \frac{\in}{2^n}\right) = \sum_{n=1}^{\infty} m^* A_n + \varepsilon$$

Since ε is arbitrary, we have

$$m^*\left(\bigcup_{n=1}^{\infty} A_n\right) \le \sum_{n=1}^{\infty} m^* A_n$$

Also

$$m_*\left(\bigcup_{n=1}^{\infty} A_n\right) \ge m_*\left(\bigcup_{n=1}^{N} A_n\right) \ge \sum_{n=1}^{N} m_* A_n - m_*\left(\bigcap_{n=1}^{N} A_n\right)$$

Since A_n, $n = 1, 2, \ldots$ are pairwise disjoint,

$$\bigcap_{n=1}^{N} A_n = \phi \quad \therefore m_*\left(\bigcap_{n=1}^{N} A_n\right) = 0$$

Letting N tend to infinity,

$$m_*\left(\bigcup_{n=1}^{\infty} A_n\right) \ge \sum_{n=1}^{\infty} m_* A_n$$

Also

$$m_* A_n = m^* A_n = mA_n \quad \text{for } n = 1, 2, 3, \ldots$$

$$\therefore \qquad \sum_{n=1}^{\infty} m A_n \leq m_*\left(\bigcup_{n=1}^{\infty} A_n\right) \leq m^*\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} m A_n$$

This implies

$$m_*\left(\bigcup_{n=1}^{\infty} A_n\right) = m^*\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} m A_n .$$

Corollary 1. If A_n, $n = 1, 2, 3, \ldots$ are measurable subsets of $[a, b]$ and if $A_i \subseteq A_{i+1}$ for all $i\,(A_{i+1} \subseteq A_i$ for all $i)$ then $\bigcup_{n=1}^{\infty} A_n \left(\bigcap_{n=1}^{\infty} A_n\right)$ is measurable and

$$m\left(\bigcup_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} m A_n \left(m\left(\bigcap_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} m A_n\right).$$

The sets A_1, $(A_2 - A_1)$, $(A_3 - A_2), \ldots, (A_{n+1} - A_n)$ are pairwise disjoint and measurable. Also

$$\bigcup_{i=1}^{n} A_n = A_1 \cup (A_2 - A_1) \cup (A_3 - A_2) \cup \ldots \cup (A_n - A_{n-1})$$

$$\therefore \; m\left(A_1 \cup \bigcup_{i=1}^{n-1} (A_{i+1} - A_i)\right) = m A_1 + \sum_{i=1}^{n-1} m\left(A_{i+1} - A_i\right)$$

$$= m A_1 + \sum_{i=1}^{n-1} \left(m A_{i+1} - m A_i\right) = m A_n$$

Letting n tend to infinity

$$m\left(\bigcup_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} m A_n$$

Now for each $n \in N$, $C(A_n)$ is measurable and $m C(A_n) = (b - a) - m A_n$.
If $A_{i+1} \subseteq A_i$, then $C\left(A_{i+1}\right) \supseteq C\left(A_i\right)$, hence

$$m\left(\bigcup_{n=1}^{\infty} C(A_n)\right) = \lim_{n\to\infty} m\, C(A_n) = b - a - \lim_{n\to\infty} m A_n$$

$\Rightarrow \qquad \bigcup_{n=1}^{\infty} C(A_n)$ is measurable

But $\qquad \bigcup_{n=1}^{\infty} C(A_n)$ is the complement of $\bigcap_{n=1}^{\infty} A_n$.

Hence, $\qquad \bigcap_{n=1}^{\infty} A_n$ is also measurable, and

$$m\left(C\left(\bigcap_{n=1}^{\infty} A_n\right)\right) = b - a - m\left(\bigcap_{n=1}^{\infty} A_n\right)$$

$$\Rightarrow \qquad m\left(\bigcap_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} m A_n$$

Corollary 2. If $A_1, A_2, \ldots, A_n$ are measurable subsets of $[a, b]$, then $\bigcup_{n=1}^{\infty} A_n$ is measurable and

$$m\left(\bigcup_{n=1}^{\infty} A_n\right) \le \sum_{n=1}^{\infty} m A_n$$

Moreover $\bigcap_{n=1}^{\infty} A_n$ is measurable.

$$\left[Hint: \bigcup_{n=1}^{\infty} A_n = A_1 \cup (A_2 - A_1) \cup (A_3 - (A_1 \cup A_2)) \ldots \text{ and } \bigcap_{n=1}^{\infty} A_n = C\left(\bigcup_{n=1}^{\infty} C(A_n)\right)\right]$$

Before proceeding to develop the main properties of measurable sets, it is convenient to remove the restriction that they should be bounded. Since the outer measure has the advantage that it is defined for all sets, though it is not countably additive, we define the set E to be *Lebesgue measurable*, (more briefly *measurable*) if for each set A we have

$$m^* A = m^*(A \cap E) + m^*(A \cap C(E))$$

Since $\qquad A = (A \cap E) \cup (A \cap C(E)),$

$$\therefore \qquad m^* A \le m^*(A \cap E) + m^*(A \cap C(E))$$

Thus, the set E will be measurable if for each set A,

$$m^* A \ge m^*(A \cap E) + m^*(A \cap C(E))$$

If E is measurable then $m^* E$ is denoted as mE.

The definition of measurability being symmetric in E and $C(E)$, hence for every measurable set its complement is measurable.

Clearly ϕ and **R** are measurable. The theorems which have already been proved for measurable sets defined on closed interval hold good for arbitrary sets, some of which are stated as under.

Theorem 5. *If E_1 and E_2 are measurable then $E_1 \cup E_2$ is measurable.* E_1 being measurable, by definition, for each set A,

$$m^* A = m^*(A \cap E_1) + m^*(A \cap C(E_1)) \qquad \ldots(1)$$

Again E_2 is measurable. Thus, taking $A \cap C(E_1)$ in place of A in the condition of measurability, we get

$$m^*(A \cap C(E_1)) = m^*(A \cap C(E_1) \cap E_2) + m^*(A \cap C(E_1) \cap C(E_2)) \qquad \ldots(2)$$

Now

$$(A \cap E_1) \cup (A \cap C(E_1) \cap E_2) = A \cap (E_1 \cup E_2) \text{ as a disjoint union}$$

$$\therefore \qquad m^*(A \cap (E_1 \cup E_2)) = m^*((A \cap E_1) \cup (A \cap C(E_1) \cap E_2))$$

$$\le m^*(A \cap E_1) + m^*(A \cap C(E_1) \cap E_2) \qquad \ldots(3)$$

(by Ex. 2)

Using (2) in (3), we get

$$m^*(A \cap (E_1 \cup E_2)) \le m^*(A \cap E_1) + m^*(A \cap C(E_1) - m^*(A \cap C(E_1) \cap C(E_2))$$

or $$m^* A \ge m^*(A \cap (E_1 \cup E_2) + m^*(A \cap C(E_1) \cap C(E_2)) \qquad \text{(from (1))}$$

Also $$m^* A \le m^*(A \cap (E_1 \cup E_2)) + m^*(A \cap C(E_1 \cup E_2))$$

Hence $$m^* A = m^*(A \cap (E_1 \cup E_2)) + m^*(A \cap C(E_1 \cup E_2))$$

Theorem 6. *If A is any set and* $E_1, E_2, \ldots, E_n$ *is a finite sequence of disjoint measurable sets, then*

$$m^*\left(A \cap \left(\bigcup_{i=1}^{n} E_i\right)\right) = \sum_{i=1}^{n} m^*(A \cap E_i).$$

[*Hint*: This can be proved by induction on n and using]

$$A \cap \left(\bigcup_{i=1}^{n} E_i\right) \cap E_n = A \cap E_n$$

and

$$A \cap \left(\bigcup_{i=1}^{n} E_i\right) \cap C(E_n) = A \cap \left(\bigcup_{i=1}^{n-1} E_i\right) \qquad [\because E_n \text{are disjoint.}]$$

Theorem 7. *For any sequence of sets* $\{E_i\}$

$$m^*\left(\bigcup_{i=1}^{\infty} E_i\right) \le \sum_{i=1}^{\infty} m^*(E_i).$$

The inequality is trivial if any one of the sets has infinite outer measure. Therefore we consider the case in which the sets have finite outer measure. If E_i's are open then measurable (by Cor. page 789), and so the result follows by Cor. 2. Theorem 4. Otherwise, ref. Ex. 7, p. 797, given $\in > 0$ $\exists$ open sets F_i such that

$$F_i \supseteq E_i, \; i = 1, 2, 3, \ldots$$

and $$mF_i = m^* F_i < m^* E_i + \in/2^i$$

$$mF_i = m^* F_i \ge m^* E_i, \; i = 1, 2, 3, \ldots$$

$$m^*\left(\bigcup_{i=1}^{\infty} E_i\right) \le m^*\left(\bigcup_{i=1}^{\infty} F_i\right) = m\left(\bigcup_{i=1}^{\infty} F_i\right) \le \sum_{i=1}^{\infty} mF_i = \sum_{i=1}^{\infty} mF_i$$

$$\le \sum_{i=1}^{\infty}\left(m^* E_i + \in/2^i\right) \le \sum_{i=1}^{\infty} m^* E_i + \in$$

(F_i being open and hence measurable)

But $\in$ being arbitrary small number,

$$\therefore \qquad m^*\left(\bigcup_{i=1}^{\infty} E_i\right) \le \sum_{i=1}^{\infty} m^* E_i.$$

Theorem 8. *The interval* $[a, \infty]$ *is measurable.*

In order to show that $[a, \infty[$ is measurable we have to show that for any set A.

$$m^* A \geq m^*(A \cap]-\infty, a[) + m^*(A \cap [a, \infty[) \quad ...(1)$$

Let $A_1 = A \cap]-\infty, a[$, and $A_2 = A \cap [a, \infty[$.

Since $m^* A = \inf_{A \subseteq \bigcup_n I_n} \sum_{n=1}^{\infty} l(I_n)$, therefore for each $\in > 0$ there exist open intervals $I_n, n \in N$, such that

$$A \subseteq \bigcup_n I_n, \text{ and } m^* A + \in > \sum_{n=1}^{\infty} l(I_n) \quad ...(2)$$

Define

$$I_n' = I_n \cap]-\infty, a[\text{ and } I_n'' = I_n \cap [a, \infty[.$$

so that

$$l(I_n) = l(I_n') + (I_n''). \qquad [\because]-\infty, a[\text{ and } [a, \infty[\text{ are disjoint}]$$

Then

$$A_1 = A \cap]-\infty, a[\subseteq \bigcup_n I_n \cap]-\infty, a[= \bigcup_n (I_n \cap]-\infty, a)$$

i.e.,

$$A_1 \subseteq \bigcup_n (I_n')$$

Similarly

$$A_2 \subseteq \bigcup_n (I_n)$$

Now $m^* A_1 + m^* A_2 \leq \sum_{n=1}^{\infty} l(I_n') + \sum_{n=1}^{\infty} l(I_n'') = \sum_{n=1}^{\infty} l(I_n) \leq m^* A + \in$ (by (*ii*))

Since $\in$ is arbitrary,

$$\therefore \qquad m^* A \geq m^* A_1 + m^* A_2$$

Hence, $[a, \infty[$ is measurable.

Corollary. Every interval is measurable. Also every open (closed) set in **R** is measurable.

2. SETS OF MEASURE ZERO

A subset A of **R** is said to be a *set of measure zero* if for any $\in > 0$ there exists a sequence of bounded open intervals $I_1, I_2, \ldots$ such that

(*i*) $A \subseteq \bigcup_{n=1}^{\infty} I_n$

(*ii*) $\sum_{n=1}^{\infty} l(I_n) \leq \in$

From the definition it is clear that every subset of a set of measure zero has measure zero. Moreover, if $A = \bigcup_{n=1}^{\infty} A_n$, where A_n has measure zero for all n, then A has measure zero.

Theorem 9. *The following statements regarding the set E are equivalent:*

(*i*) E is measurable.

(*ii*) *For all* $\varepsilon > 0, \exists\, O$ *– an open set,* $O \supseteq E$ *such that* $m^*(O - E) \le \varepsilon$.

(*iii*) $\exists\, G$, a $G_{\delta-set}$, $G \supseteq E$ *such that* $m^*(G - E) = 0$, (*A set G is said to be* G_δ *if* $G = \bigcap_{i=1}^{\infty} G_i$, *each* G_i *is an open set.*)

(*iv*) *For all* $\varepsilon > 0, \exists\, F$ *– a closed set,* $F \subseteq E$, *such that* $m^*(E - F) \le \varepsilon$.

(*v*) $\exists\, F$, a F_σ-set, $F \subseteq E$ *such that* $m^*(E - F) = 0$. (*A set F is said to be* F_σ *if* $F = \bigcup_{i=1}^{\infty} F_i$, *each* F_i *is a closed set.*)

(*i*) ⇒ (*ii*). Let us first consider the case when $mE < \infty$. E being measurable, $\exists$ an open set $O \supseteq E$ such that

$$m^* O \le m^* E + \varepsilon$$

$$\therefore \quad m^*(O - E) = m^* O - m^* E \le \varepsilon$$

If $mE = \infty$, let $E \subseteq \bigcup_{n=1}^{\infty} I_n$ where I_n's are disjoint finite intervals. Define $E_n = E \cap I_n$, then $E = \bigcup_{n=1}^{\infty} E_n$. For each n, E_n is measurable, belong to intersection of two measurable sets. Moreover $E_n \subseteq I_n$ and $m(I_n) < \infty$. Hence $mE_n < \infty$. From the above result, it follows that, for each $\in > 0$ $\exists$ an open set O_n such that $O_n \supseteq E_n$ and $m(O_n - E_n) \le \frac{\in}{2^n}$, for $n = 1, 2, 3, \ldots$

Write $O = \bigcup_{n=1}^{\infty} O_n$, then O is an open set. Consider

$$O - E = \bigcup_{n=1}^{\infty} O_n - \bigcup_{n-1}^{\infty} E_n \subseteq \bigcup_{n=1}^{\infty} (O_n - E_n)$$

$$m^*(O - E) \le m^*\left(\bigcup_{n=1}^{\infty} (O_n - E_n)\right) \le \sum_{n=1}^{\infty} m^*(O_n - E_n) \le \sum_{n=1}^{\infty} \frac{\in}{2^n} = \in.$$

$\therefore$ (*i*) ⇒ (*ii*)

(*ii*) ⇒ (*iii*). For each n, $\exists$ an open set $O_n \supseteq E$ such that $m^*(O_n - E) < \frac{1}{n}$.

Define $G = \bigcap_{n=1}^{\infty} O_n$; then G is a G_δ-set. Since $E \subseteq O_n$ for each n, so $E \subseteq \bigcap_{n=1}^{\infty} O_n = G$. Thus,

$$m^*(G - E) \le m^*(O_n - E) < \frac{1}{n}, \text{ for each } n,$$

$\therefore \qquad m^*(G - E) = 0$.

(*iii*) $\Rightarrow$ (*i*). Let G be a G_δ-set containing E such that $m^*(G - E) = 0$ by Ex. 1 (page 794), $G - E$ is measurable. G being a G_δ-set is measurable.

Hence $E = G - (G - E)$ is measurable.

Similarly equivalence in the remaining statements can be established.

If $m^* E$ is finite the above statements are equivalent to the following:

(*v*) Given $\varepsilon > 0$, there is a disjoint finite union $\bigcup_{i=1}^{n} I_i$ of open intervals such that $m^*\left(\bigcup_{i=1}^{n} I_i \,\Delta\, E\right) < \varepsilon$,

where $\bigcup_{i=1}^{n} I_i \,\Delta\, E = \left(\bigcup_{i=1}^{n} I_i - E\right) \cup \left(E - \bigcup_{i=1}^{n} I_i\right)$.

The intervals I_i can be either open, closed or half open but we will give the proof for open intervals.

Suppose E is measurable then by (*ii*) $\forall\, \varepsilon > 0, \exists$ an open set $O, O \supseteq E$ such that

$$m(O - E) < \frac{\varepsilon}{2} \qquad \text{...(1)}$$

mE being finite, so mO is finite, O being the union of disjoint open intervals, $I_i, i = 1, 2, 3, \ldots \infty$, such that

$$m^*\left(\bigcup_{i=1}^{\infty} I_i\right) \le \sum_{i=1}^{\infty} m^* I_i < \infty$$

Hence, $\exists$ a positive integer n such that

$$\sum_{i=n+1}^{\infty} m I_i < \varepsilon,\ i.e., m\left(\bigcup_{i=n+1}^{\infty} I_i\right) < \frac{\varepsilon}{2}$$

[by Cauchy's General principle of convergence for series]

$$\Rightarrow \qquad m\left(O - \bigcup_{i=1}^{n} I_i\right) < \frac{\varepsilon}{2} \qquad \text{...(2)}$$

Writing $U = \bigcup_{i=1}^{n} I_i$,

$$E \,\Delta U = \left(E - \bigcup_{i=1}^{n} I_i\right) \cup \left(\bigcup_{i=1}^{n} I_i - E\right) \subseteq \left(O - \bigcup_{i=1} I_i\right) \cup (O - E)$$

$$\therefore \qquad m^*(E\ \Delta U) \le m^*\left(\left(O - \bigcup_{i=1}^{n} I_i\right) \cup (O - E)\right)$$

$$\le m^*\left(O - \bigcup_{i=1}^{n} I_i\right) + m^*(O - E) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \qquad \text{[from (1) and (2)]}$$

Conversely, for every given set E and $\varepsilon > 0, \exists$ an open set O containing E such that $m^* O \leq m^* E + \varepsilon$. Now we show that $m^* (O - E) < \varepsilon$ whenever

$$m^* \left(E \,\Delta \bigcup_{i=1}^{n} I_i \right) < \frac{\varepsilon}{3}$$

Let $U = \bigcup_{i=1}^{n} (I_i \cap O)$, where I_i, $i = 1, 2, ..., n$, are given open intervals corresponding to the same ε.

$$U \subseteq \bigcup_{i=1}^{n} I_i$$

Thus, $$U \,\Delta E \subseteq (E - U) \cup \left(\bigcup_{i=1}^{n} I_i - E \right)$$

But $$E - U = E - \left(\left(\bigcup_{i=1}^{n} I_i \right) \cap O \right)$$

or $$E - U = E \cap C(O) \cup E \cap C \left(\bigcup_{i=1}^{n} I_i \right) = E - \bigcup_{i=1}^{n} I_i$$

As $E \subseteq O$ $\quad \therefore \; E \cap C(O) = \phi$

$\therefore$ $$U \,\Delta E \subseteq \left(E - \bigcup_{i=1}^{n} I_i \right) \cup \left(\bigcup_{i=1}^{n} I_i - E \right) = E \Delta \left(\bigcup_{i=1}^{n} I_i \right)$$

$\Rightarrow$ $$m^* (U \,\Delta E) \leq m^* \left(E \Delta \bigcup_{i=1}^{\infty} I_i \right) < \frac{\varepsilon}{3} \qquad ...(3)$$

Now $\quad E \subseteq U \cup (U \,\Delta E)$

$\therefore$ $$m^* E \leq m^* U + m^* (U \,\Delta E) < mU + \frac{\varepsilon}{3} \qquad \text{[using (3)]}$$

or $$m^* E < mU + \frac{\varepsilon}{3} \qquad ...(4)$$

Also $\quad O - E \subseteq (O - U) \cup (U \,\Delta E)$

$$m^* (O - E) \leq m^* (O - U) + m^* (U \,\Delta E) < m(O - U) + \frac{\varepsilon}{3} \qquad \text{[using (3)]}$$

or $$m^* (O - E) < mO - mU + \frac{\varepsilon}{3}$$

Also O and U have finite measure

$\therefore$ $$m^* (O - E) < m^* E - mU + \frac{2\varepsilon}{3} \qquad \left[\because m^* E + \frac{\varepsilon}{3} > m\,O \right]$$

$$\therefore \qquad m^*(O - E) < \frac{\varepsilon}{3} + \frac{2\varepsilon}{3} \qquad \text{[using (4)]}$$

$\Rightarrow$ E is measurable

Remark: The result (*vi*) indicates that a measurable set with finite measure may be approximated by a finite union of disjoint open intervals. It must be noted that the union of these intervals in general neither contains nor is contained in the set which we are approximating.

3. BOREL SETS

A class **F** of subsets of an arbitrary field F is said to be σ-*Algebra* (sigma algebra) if

(*i*) $F \in \mathbf{F}$

(*ii*) $F_1, F_2, \ldots F_n, \ldots \in \mathbf{F}$

implies $\bigcup_{l=1}^{\infty} F_l \in \mathbf{F}$ and $F_1 - F_2 \in \mathbf{F}$ whenever, $F_1, F_2 \in \mathbf{F}$.

From (*i*) and (*ii*), it follows that the class **F** is closed under countable intersection. The σ-algebra generated by any class of sets is the smallest σ-algebra containing this class of sets.

The *Borel Sets* in R are the members of the σ-algebra generated by the class of intervals of the form $[a, b[$.

The Lebesgue measurable sets form a σ- algebra containing intervals because any countable union or countable intersection of measurable sets is measurable and $A_1 \cap C(A_2)$ is measurable whenever A_1 and A_2 are measurable subsets of **R**. As Borel sets in **R** being the members of the smallest σ-algebra containing the intervals of the form $[a, b[$ are themselves measurable. It is to be noted that every Lebesgue measurable set need not be a Borel set.

4. NON-MEASURABLE SETS

So far we have talked about measurable sets. a natural question arises: do there exist non-measurable sets as well? Utilising the *axiom of choice* (if $\{F_\alpha : \alpha \in A\}$ is a non-empty collection of non-empty disjoint subsets of a set X then there exists a set $E \subseteq X$ containing just one element from each set F_α), one can construct a *non-measurable set* as follows:

Divide the set of all points on the circumference of the unit circle into classes, where two points lie in the same class if and only if the arc joining them has rational length. These classes will be disjoint. By the axiom of choice there exists a set E on the circumference which has exactly one point in common with each class. The set E constructed in this manner will be such that the arc joining any two of its members will be of irrational length. The set E is *non-measurable*. To show this, displace E along the circumference by the sets of the form $E + x$, where x is any rational number. Then the entire circumference will be covered, because every point p of the circumference lies at some rational arc length say x from some point of E. Thus p belongs to $E + x$. These displacements of E are disjoint for otherwise if p belongs to $E + x$ as well as $E + y$ then p will be at rational distance x from some point of E say e_1 which

belongs to one of the equivalence class say P_1, and at a rational distance y from some point of E say e_2 which belongs to P_2 $(e_1 \neq e_2)$. This implies that p belongs to both the classes P_1 and P_2 which is not true as P_1 and P_2 are disjoint equivalence classes. Now we will show that the set is non-measurable. If the set E was measurable, then its displacements would also be measurable and they would have the same measure (see Example 5). If this measure was zero then the measure of the circumference of the given unit circle will be zero, because the circumference has been expressed as a countable disjoint union of these displacements. Hence measure of E cannot be zero. If this measure was positive then the circumference would have infinite measure which again cannot be possible. Thus E is non-measurable.

Example 1. If $m^* E = 0$, then E is measurable.

■ Let A be any set, then $A \cap E \subseteq E$, so $0 \leq m^*(A \cap E) \leq m^* E = 0$.

$\therefore$ $$m^*(A \cap E) = 0$$

Also $$A \cap E \subseteq E, \text{ so}, 0 \leq m^*(A \cap E) \leq m^* E = 0$$

$$A \supseteq A \cap C(E) \qquad \therefore\ m^* A \geq m^*(A \cap C(E))$$

Now

$$m^*(A \cap C(E)) + m^*(A \cap E) = m^*(A \cap C(E)) \leq m^* A \qquad ...(1)$$

But

$$A = (A \cap E) \cup (A \cap C(E))$$

$\therefore$ $$m^* A \leq m^*(A \cap E) + m^*(A \cap C(E)) = m^*(A \cap C(E)) \qquad ...(2)$$

From equations (1) and (2), we get

$$m^* A = m^*(A \cap E) + m^*(A \cap C(E))$$

Hence, E is measurable.

Example 2. A set consisting of one point is measurable and its measure is zero.

■ Let $\{x\}$ be the given singleton set. Since $x \in I_n = \left]x - \frac{1}{2^{n+1}}, x + \frac{1}{2^{n+1}}\right[$ for all n and $l(I_n) = \frac{1}{2^n}$.

$\therefore$ $$m^*\{x\} = \inf_{\{x\} \subset I_n} l(I_n) = 0$$

$\Rightarrow$ $\{x\}$ is measurable.

Example 3. A countable set is measurable and its measure is zero.

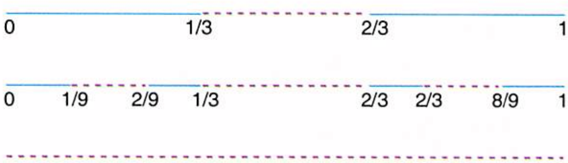
■ Let $A = \bigcup_{i=1}^{\infty} \{x_i\}$ be a given countable set, then

$$m^* A = m^*\left(\bigcup_{i=1}^{\infty} \{x_i\}\right) \leq \sum_{i=1}^{\infty} m^*\{x_i\} = 0 \qquad \text{[by Example 2]}$$

Note: From the above example, [0, 1] is not countable, since $m^*[0, 1] = 1 \neq 0$, however, one should not conclude that the sets of measure zero consist only of at most countable number of points. Indeed we have the following example.

Example 4. There exist uncountable sets of measure zero.

■ Let us construct the Cantor sets by dividing the interval [0, 1] into three equal parts by the points $\frac{1}{3}, \frac{2}{3}$ and remove open interval $\left]\frac{1}{3}, \frac{2}{3}\right[$ from [0, 1]. Then dividing each of the remaining closed intervals $\left[0, \frac{1}{3}\right]; \left[\frac{2}{3}, 1\right]$ into three parts by the point $\frac{1}{9}, \frac{2}{9}$ for the first interval $\left[0, \frac{1}{3}\right]$ and by the points $\frac{7}{9}, \frac{8}{9}$ for the second interval $\left[\frac{2}{3}, 1\right]$ and remove the middle open intervals $\left]\frac{1}{9}, \frac{2}{9}\right[, \left]\frac{7}{9}, \frac{8}{9}\right[$. Next divide each of the remaining four intervals into three equal parts and remove the middle intervals from these. Continue this process indefinitely. By this process we remove an open set G_0 from [0, 1], the set G_0 is the union of a countable family of open intervals, i.e.,



$$G_0 = \left]\frac{1}{3}, \frac{2}{3}\right[\cup \left]\frac{1}{9}, \frac{2}{9}\right[\cup \left]\frac{7}{9}, \frac{8}{9}\right[\ldots$$

The complement of G_0 in [0, 1] is denoted by P_0. Clearly the set P_0 is the set of all those points which are left after deleting the intervals, i.e., P_0 is the intersection of all the closed intervals which are left after each iteration.

The set P_0 is *an uncountable set*. Let, if possible, P_0 be countable. We enumerate the points of P_0 by x_i, $i = 1, 2, \ldots \infty$.. Let the ternary representation (i.e., expansion in the scale of three of the form $\sum_{k=1}^{\infty} \frac{a_k}{3^k}$, where each a_k is either zero or two; see also the note) of the points of P_0 be denoted by

$$x_i = 0 a_{i1}\, a_{i2}\, a_{i3} \ldots, \; i = 1, 2, 3, \ldots$$

where

$$a_{ij} = 0 \text{ or } 2.$$

Now we construct a point

$$y = 0.b_1 b_2 b_3 \ldots$$

where

$$b_1 = 0 \text{ if } a_{ii} = 2$$

$$i = 1, 2, 3, \ldots$$

$$= 2 \text{ if } a_{ii} = 0$$

Since b_i's are either zero or 2, therefore the new point y constructed as above which is different from x_i for any i, belongs to P_0 which contradicts the fact that the set P_0 is countable. Hence the set P_0 is uncountable. Now

$$m(G_0) = \frac{1}{3} + \frac{2}{9} + \frac{4}{27} + \cdots = \frac{\frac{1}{3}}{1 - \frac{2}{3}} = 1$$

Hence, the uncountable set P_0 has measure

$$m(P_0) = m([0, 1] - G_0) = 0$$

The sets G_0 and P_0 are called *Cantor Sets.*

Note: Let $x = \frac{7}{9} \in P_0$. Then by the ternary representation of x (*i.e.* expansion of x in the scale of 3) we mean

$$\frac{7}{9} = \frac{2.3 + 1}{3^2} = \frac{2}{3} + \frac{1}{3^2} = .21 = .202222\ldots = \frac{2}{3} + \frac{0}{3^2} + \frac{2}{3^3} + \frac{2}{3^4} + \cdots$$

Similarly,

$$\frac{8}{9} = \frac{2.3 + 2}{9} = \frac{2}{3} + \frac{2}{3^2} = .22 = .22000\ldots$$

Example 5. Show that for any set A, $m^* A = m^* (A + x)$, where

$$A + x = \{y + x : y \in A\}$$

- Let $A \subseteq \bigcup_{n=1}^{\infty} I_n$; then for each $\varepsilon > 0$, we have

$$m^* A + \varepsilon \geq \sum_{n=1}^{\infty} l(I_n)$$

Also $A + x \subseteq \bigcup_{n=1}^{\infty} (I_n + x)$. Thus for each $\varepsilon > 0$,

$$m^* (A + x) \leq \sum_{n=1}^{\infty} l(I_n + x) = \sum_{n=1}^{\infty} l(I_n) \leq m^* A + \varepsilon$$

ε being arbitrary,

$$m^* (A + x) \leq m^* A \qquad \ldots(1)$$

Writing $A = A + x - x$, from (1), it follows that

$$m^* A = m^* ((A + x) - x) \leq m^* (A + x) \qquad \ldots(2)$$

From (1) and (2), we get the desired result.

Example 6. If A is measurable, then for each x the set $A + x = \{y + x : y \in A\}$ is measurable, and the measures are the same.

- As A is measurable, by Theorem 9, for each $\in > 0$, $\exists$ an open set O, $O \supseteq A$ such that $m^* (O - A) < \in$. The set $O + x$ is open and $O + x \supseteq A + x$.

Moreover $(O+x)-(A+x)=(O-A)+x$

and

$$m^*(O-A)=m^*((O-A)+x)<\in$$

$\Rightarrow$ $$m^*(O+x-(A+x))<\in$$

Hence $A+x$ is measurable

Example 7. Show that, for any set A and any $\in>0, \exists$ an open set O such that $A\subseteq O$ and $m^* O\le m^* A+\in$

■ Let I_n, $n=1, 2, 3, \ldots$ be open intervals such that

$$A\subseteq \bigcup_{n=1}^{\infty} I_n$$

By definition,

$$m^* A=\inf_{A\subseteq \bigcup_n I_n} \sum_{n=1}^{\infty} l(I_n)$$

Hence given $\in>0, \exists$ intervals I_n, $n=1, 2, 3, \ldots$ such that

$$m^* A\ge \sum_{n=1}^{\infty} l(I_n)-\frac{\in}{2}$$

If $I_n=[a_n, b_n[$, and $I'_n=]a_n-\in/2^{n+1}, b_n+\in/2^{n+1}[$ so that $A\subseteq \bigcup_{n=1}^{\infty} I'_n$, and if $O=\bigcup_{n=1}^{\infty} I'_n$, then O is open set.

$$m^* O\le \sum_{n=1}^{\infty} l(I'_n)=\sum_{n=1}^{\infty} l(I_n)+\frac{1}{2}\in\le m^* A+\frac{1}{2}\in+\frac{1}{2}\in$$

$\therefore$ $$m^* O\le m^* A+\in.$$

5. MEASURABLE FUNCTIONS

As already discussed in the chapter on Riemann integration that if a function f is either continuous or the set of limit points of discontinuities of f is finite then f is integrable. We shall see that a function may be Lebesgue integrable on $[a, b]$ even if this might not be the case. In fact, a much less restrictive condition than continuity viz. measurability is needed to ensure integrability of f on $[a, b]$. The definition of measurability of functions applies to both bounded as well as unbounded functions.

Definition. Let f be a function defined on $[a, b]$. We call f to be a *measurable function* if for each $\alpha\in \mathbf{R}$, the set $\{x : f(x)>\alpha\}$ is a *measurable set, i.e.*, f is a measurable function if for every real number α, the inverse image of $]\alpha, \infty[$ is a measurable set. As $]\alpha, \infty[$ is an open set and if f is continuous, then the inverse image under f of $]\alpha, \infty[$ is open. Open sets being measurable, hence every continuous function is measurable.

Theorem 10. *The function f on $[a, b]$ is measurable if and only if any one of the following conditions hold:*

(i) $\{x: f(x) > \alpha\}$ *is measurable set for every real* α

(ii) $\{x: f(x) \geq \alpha\}$ *is measurable set for every real* α

(iii) $\{x: f(x) < \alpha\}$ *is measurable set for every real* α

(iv) $\{x: f(x) \leq \alpha\}$ *is measurable set for every real* α

Suppose f is measurable, then by definition, the set $\{x: f(x) > \alpha\}$ for every $\alpha \in \mathbf{R}$ is a measurable set. The set $\{x: f(x) \leq \alpha\}$ is the complement of the set $\{x: f(x) > \alpha\}$ in the space of all measurable sets, and the complement of a measurable set is measurable. Therefore the measurability of the set $\{x: f(x) > \alpha\}$ implies the measurability of the set $\{x: f(x) \leq \alpha\}$ *i.e.*, $(i) \Rightarrow (iv)$.

Again $(i) \Rightarrow (ii)$. Since f is measurable, and $\alpha \in \mathbf{R}$ then each of the sets $\{x : f(x) > \alpha - 1/n\}$, $n = 1, 2, \ldots$ is measurable and $\{x : f(x) \geq \alpha\} = \bigcap_{n=1}^{\infty} \{x : f(x) > \alpha - 1/n\}$. This being the arbitrary intersection of measurable sets is measurable.

$(i) \Rightarrow (iii)$. Because the set $\{x : f(x) < \alpha\}$ is the complement of the set $\{x : f(x) \geq \alpha\}$ and the set $\{x : f(x) \geq \alpha\} = \bigcap_{n=1}^{\infty} \{x : f(x) > \alpha - 1/n\}$. Therefore measurability of (i) implies measurability of (iii).

$(iii) \Rightarrow (iv)$. Since the set $\{x : f(x) \leq \alpha\} = \bigcap_{n=1}^{\infty} \{x : f(x) < \alpha + 1/n\}$ and each of the sets $\{x : f(x) < \alpha + 1/n\}$ is measurable. The set $\{x : f(x) > \alpha\}$ being the complement of the set $\{x : f(x) \leq \alpha\}$ and $\{x : f(x) \leq \alpha\}$ is measurable because (iii) holds.

$(iii) \Rightarrow (i)$. Now for each n, the set $\{x : f(x) > \alpha - 1/n\}$ is the complement of the set $\{x : f(x) \leq \alpha - 1/n\}$, and $\{x : f(x) \geq \alpha\} = \bigcap_{n=1}^{\infty} \{x : f(x) > \alpha - 1/n\}$ is the complement of

$$\bigcup_{n=1}^{\infty} \{x : f(x) \leq \alpha - 1/n\} = \{x : f(x) < \alpha\}.$$

Hence any of these conditions may be used to define measurability of functions.

Theorem 11. *If f is measurable then $|f|$ is measurable.*

Since the set $\{x : |f(x)| < \alpha\} = \{x : f(x) < \alpha\} \cap \{x : f(x) > -\alpha\}$ and each of the sets

$\{x : f(x) < \alpha\}$, $\{x : f(x) > -\alpha\}$ is measurable. Hence $|f|$ is measurable.

Theorem 12. *If f is a measurable function on $[a, b]$ and $k \in \mathbf{R}$, then $f + k$ and kf are measurable.*

If $\alpha \in \mathbf{R}$, then $\{x : f(x) + k > \alpha\} = \{x : f(x) > \alpha - k\}$. The set on the right is measurable because f is measurable. Hence the set on the left is measurable.

If $k > 0$, $k f(x) > \alpha \Rightarrow f(x) > \dfrac{\alpha}{k}$.

Hence the set $\{x : k f(x) > \alpha\} = \left\{x : f(x) > \dfrac{\alpha}{k}\right\}$ ...(1)

and if $k < 0, kf(x) > \alpha \Rightarrow f(x) < \dfrac{\alpha}{k}$. And the set

$$\{x : kf(x) > \alpha\} = \left\{x : f(x) < \frac{\alpha}{k}\right\} \qquad \text{...(2)}$$

Each of the sets on the right of (1) and (2) is measurable. Hence the set on the left is measurable. Clearly if f is measurable the $-f$ is measurable by taking $k = -1$.

Example 8. Show that if f is measurable then the set $\{x : f(x) = \alpha\}$ is measurable for each extended real number α.

■ For α to be finite, $\alpha \in \mathbf{R}$ the set

$$\{x : f(x) = \alpha\} = \{x : f(x) \geq \alpha\} \cap \{x : f(x) \leq \alpha\}$$

and each of these sets is measurable and so the set $\{x : f(x) = \alpha\}$.

For $\alpha = +\infty$, the set $\{x : f(x) = \infty\} = \bigcap_{n=1}^{\infty} \{x : f(x) > n\}$ is measurable.

Similarly, for $\alpha = -\infty$.

Example 9. Constant functions are measurable.

■ Given $f(x) = C$ for all $x \in \mathbf{R}$.

Let $\alpha \in \mathbf{R}$, if $C \geq \alpha$, then the set $\{x : f(x) \geq \alpha\}$ is $\mathbf{R}$ which is measurable.

And if $C < \alpha$, then the set $\{x : f(x) \geq \alpha\}$ is ϕ, which is again measurable, α being arbitrary. Hence f is measurable.

6. MEASURABILITY OF THE SUM, DIFFERENCE, PRODUCT AND QUOTIENT OF MEASURABLE FUNCTIONS

Theorem 13. *If f_1 and f_2 are measurable on $[a, b]$ then so are $f_1 + f_2$, $f_1 - f_2$, $f_1 f_2$ and f_1/f_2 provided $f_2 \neq 0$ on $[a, b]$.*

Let $q_1, q_2, \ldots, q_n \ldots$ be an enumeration of the set of rational numbers. Let $x \in [a, b]$ and $\alpha \in \mathbf{R}$.

$f_1(x) > \alpha - f_2(x)$ if and only if there is a rational number q_n such that

$$f_1(x) > q_n > \alpha - f_2(x)$$

Hence, the set

$$\{x : f_1(x) + f_2(x) > \alpha\}$$

$$= \bigcup_{n=1}^{\infty} \left[\{x : f_1(x) > q_n\} \cap \{x : \alpha - f_2(x) < q_n\}\right] \qquad \text{...(1)}$$

CHAPTER 20

For any $n \in \mathbf{N}$ the set $\{x : f_1(x) > q_n\}$ is measurable, since f_1 is a measurable function. And the set $\{x : \alpha - f_2(x) < q_n\} = \{x : f_2(x) > \alpha - q_n\}$ is measurable because f_2 is measurable. The set on the right of (*i*) is measurable which establishes that the sum $f_1(x) + f_2(x)$ is measurable.

Again $f_1(x) - f_2(x) = f_1(x) + (-f_2(x))$ and hence measurable. Now to show that $f_1(x)\cdot f_2(x)$ is measurable, we will first show that the square of a measurable function is measurable. If g is a measurable function on $[a, b]$ and $\alpha \geq 0$, then

$$\{x : (g(x))^2 > \alpha\} = \{x : g(x) > \sqrt{\alpha}\} \cup \{x : g(x) < -\sqrt{\alpha}\}$$

Each of the sets on the right is measurable and hence their union is measurable.

If $\alpha < 0$ then the set $\{x : (g(x))^2 > \alpha\}$ is in the interval $[a, b]$, hence measurable.

Now f_1 and f_2 are measurable then $(f_1 + f_2)^2$ and $(f_1 - f_2)^2$ are also measurable.

Since $f_1(x) f_2(x) = \dfrac{1}{4}\left[(f_1(x) + f_2(x))^2 - (f_1(x) - f_2(x))^2\right]$, it follows that $f_1(x)\cdot f_2(x)$ is measurable.

At last to show that $\dfrac{f_1(x)}{f_2(x)}$, $f_2(x) \neq 0$ on $[a, b]$ is measurable, we will first show that $\dfrac{1}{f_2(x)}$, $f_2(x) \neq 0$ on $[a, b]$ is measurable.

As $f_2(x) \neq 0$ on $[a, b]$ we have

$$\left\{x : \frac{1}{(x)} > \alpha\right\}$$

$$= \begin{cases} \{x : f_2(x) > 0\} & \text{if } \alpha = 0 \\ \{x : f_2(x) > 0\} \cap \{x : f_2(x) < 1/\alpha\} & \text{if } \alpha > 0 \\ \{x : f_2(x) > 0\} \cup \{x : f_2(x) < 0\} \cap \{x : f_2(x) < 1/\alpha\} & \text{if } \alpha < 0 \end{cases}$$

Each of the sets on the right is measurable and if f_2 is measurable on $[a, b]$ then $1/f_2$, $f_2(x) \neq 0\ \forall\ x$ on $[a, b]$ is measurable.

Now $\dfrac{f_1(x)}{f_2(x)} = f_1(x)\cdot\dfrac{1}{f_2(x)}$ provided $f_2(x) \neq 0, \forall\ x \in [a, b]$. Hence from the previous result it follows that $\dfrac{f_1(x)}{f_2(x)}$ is measurable, when $f_2(x) \neq 0$ on $[a, b]$.

Theorem 14. *If $\{f_n\}$ is a sequence of measurable functions on $[a, b]$ such that the sequence $\{f_n(x)\}$ is bounded for every $x \in [a, b]$, then the functions*

$$G(x) = l.u.b\,\{f_1(x),\ f_2(x),\ f_3(x)\ldots\}$$

$$g(x) = g.l.b\,\{f_1(x),\ f_2(x),\ f_3(x)\ldots\}$$

$$H(x) = \limsup_{n\to\infty} f_n(x)$$

$$h(x) = \liminf_{n\to\infty} f_n(x)$$

are measurable.

Let $\alpha \in \mathbf{R}$. Let $x \in [a, b]$ be such that $G(x) > \alpha$, then $\exists$ some n for which $f_n(x) > \alpha$. Also if $f_n(x) > \alpha$, then $\underset{n}{\text{l.u.b}}\ f_n(x) > \alpha$ *i.e.*, $G(x) > \alpha$. Thus $G(x) > \alpha$ iff $f_n(x) > \alpha$ for some n.

Now consider the set

$$\{x : G(x) > \alpha\} = \bigcup_n \{x : f_n(x) > \alpha\}$$

Each f_n being a measurable function, the set $\{x : f_n(x) > \alpha\}$ is measurable, and hence their countable union $\bigcup_n \{x : f_n(x) > \alpha\}$ is measurable $\Rightarrow$ $\{x : G(x) > \alpha\}$ is measurable, α being arbitrary. Hence, the function G is measurable.

Similarly from the relation

$$\{x : g(x) < \alpha\} = \bigcup_n \{x : f_n(x) < \alpha\}$$

it follows that the function g is measurable.

Now for each $n \in \mathbf{N}$ we define the functions $G_n(x), g_n(x)$ as follows:

$$G_n(x) = \text{l.u.b}\{f_n(x),\ f_{n+1}(x),\ f_{n+2}(x),\ \ldots\}$$

$$g_n(x) = \text{g.l.b}\{f_n(x),\ f_{n+1}(x),\ f_{n+2}(x),\ \ldots\}$$

The functions $G_n(x)$ and $g_n(x)$ being the l.u.b. and g.l.b of measurable functions $f_n(x)$ are also measurable.

Also for each $x \in [a, b]$,

$$G_1(x) \geq G_2(x) \geq G_3(x) \geq \ldots G_n(x) \geq \ldots$$

and

$$g_1(x) \leq g_2(x) \leq g_3(x) \ldots \leq g_n(x) \leq \ldots$$

Now

$$H(x) = \limsup_{n\to\infty} f_n(x) = \lim_{n\to\infty} G_n(x)$$

and

$$h(x) = \liminf_{n\to\infty} f_n(x) = \lim_{n\to\infty} g_n(x)$$

The sequences $\{G_n(x)\}$ and $\{g_n(x)\}$ being monotonically decreasing and monotonically increasing, respectively, it follows that $H(x) < G_n(x)$ and $h(x) > g_n(x)$ for all n.

Now for each $\alpha \in \mathbf{R}$ the sets

$$\{x : H(x) < \alpha\} = \bigcup_n \{x : G_n(x) < \alpha\}$$

and

$$\{x : h(x) > \alpha\} = \bigcup_n \{x : g_n(x) > \alpha\}$$

are measurable. Hence the functions $H(x)$ and $h(x)$ are measurable.

Definition 1. A function f is a *Borel measurable* or a *Borel function* if for each $\alpha \in \mathbf{R}$ the set $\{x : f(x) > \alpha\}$ is a *Borel set.*

Definition 2. A property which holds everywhere except on a set of measure zero is said to hold *almost everywhere*, abbreviated as a.e.

Definition 3. Any two functions f and g are said to be *equal almost everywhere* if the set $A = \{x: f(x) \neq g(x)\}$ is of measure zero, *i.e.*, $f = g$ a.e. if $mA = 0$.

Definition 4 . A sequence $\{f_n\}$ of measurable functions defined and finite a.e. on a measurable set E is said to *converge in measure* to a measurable function f which is finite a.e. if $\underset{n\to\infty}{\text{Lt}}\ m\left(\{x : |f_n - f| \geq \alpha\}\right) = 0$, for all positive real numbers α.

Note: Convergence in measure is essentially weaker than convergence a.e., *i.e.*, there exist sequences of measurable (even continuous) functions which are divergent at every point, but which are convergent in measure.

Theorem 15. *If $f = g$ a.e. and f is a measurable function, then g is also measurable.*

Let α be a positive real number.

The set $\{x : f(x) - g(x) > \alpha\} = \{x : f(x) \neq g(x),\ f(x) - g(x) > \alpha\}$ is contained in the sets $\{x : f(x) \neq g(x)\}$ which is of measure zero. Hence $\{x : f(x) - g(x) > \alpha\}$ is of measure zero and hence measurable.

Let $\alpha < 0$, then the set

$$\{x : f(x) - g(x) > \alpha\} = \{x : f(x) = g(x)\}$$
$$\bigcup \{x : f(x) \neq g(x),\ f(x) - g(x) > \alpha\}$$

The two sets on the right hand side are measurable, the first set is the complement of the set of measure zero and the second is contained in the set of measure zero. Thus the set $(x : f(x) - g(x) > \alpha\}$ is measurable for arbitrary real α.

Hence, the function $f - g$ is measurable.

Now $g = f - (f - g)$, f and $f - g$ being measurable implies that g is measurable.

Theorem 16. *Every continuous function is measurable.*

Let $\alpha \in \mathbf{R}$ then the set $\{x : f(x) > \alpha\}$, being the inverse image of open interval $]\alpha, \infty[$ under a continuous function f is also open, and hence measurable.

The converse of this theorem may or may not be true, *i.e.*, every measurable function may or may not be continuous.

Still then every measurable function which is finite a.e. can be approximated by some continuous function in the following sense.

Theorem 17. ***N. N. Lusin.*** *Given f to be a measurable function which is finite a.e. on [a, b]. Then for every $\delta > 0$ there exists a continuous function g such that $m\{x : f \neq g\} < \delta$. Moreover if f is bounded by a constant M then g is also bounded by the same constant.*

(The proof is beyond the scope of the present chapter. Interested reader is advised to see *Theory of Functions of Real Variables,* Vol. I by I. P. Natanson).

Theorem 18. ***Egoroff.*** *If $\{f_n\}$ is a sequence of measurable functions which converge to a real valued function f a.e. on a measurable set E of finite measure, then given $\eta > 0$, there is a subset $A \subseteq E$ with $m(E - A) < \eta$ such that $\{f_n\}$ converges to f uniformly on A.*

The sequence $\{f_n\}$ being pointwise convergent to f on E a.e., let $F \subseteq E$ be the set of all $x \in E$ for which $\{f_n(x)\}$ converges to $f(x)$. Then

$$m(E - F) = 0 \qquad \text{...(1)}$$

Define the sets

$$F_{nm} = \left\{x : |f_N - f| < \frac{1}{n},\ \forall\ N \geq m\right\}$$

For a fixed n, F_{nm}'s from an increasing sequence of subsets of E (*i.e.*, $F_{n1} \subseteq F_{n2} \subseteq F_{n3} \ldots$) and $F = \bigcup_{m=1}^{\infty} E_{nm}$. The sets F_{nm} being measurable, we have $mF = \underset{m \to \infty}{\text{Lt}} F_{nm}$.

Thus given $\eta > 0$ and for each n, $\exists$ a positive integer m_n depending on η such that

$$m\left(F - F_{nm_n}\right) < \frac{\eta}{2^n} \qquad \text{...(2)}$$

If we define $A = \bigcap_{n=1}^{\infty} F_{nm_n}$, then for each $x \in E$, we have

$$|f_n - f| < \frac{1}{n} \quad \forall\ n \geq m_n$$

Clearly m_n does not depend on the choice of x. Hence $\{f_N\}$ converges to f uniformly on A.

Now we have to show that $m(E - A) < \eta$.

Consider $F - A$, where $A = \bigcap_{n=1}^{\infty} F_{nm_n}$. Thus

$$F - A \subseteq \bigcup_n \left(F - F_{nm_n}\right)$$

and

$$m(F - A) \leq \sum_{n=1}^{\infty} m\left(F - F_{nm_n}\right) < \sum_{n=1}^{\infty} \frac{\eta}{2^n} = \eta \qquad \text{(by (2))}$$

$$\therefore \quad m(E - A) = m(E - F + F - A) \leq m(E - F) + m(F - A) < \eta$$

i.e., $\quad m(E - A) < \eta$

CHAPTER 20

So far we have discussed about the measurability of the functions. A natural question arises, do there exist non-measurable functions? The answer is in the affirmative. For the function f to be *measurable on an arbitrary set* E we must have the set E on which f is defined to be measurable as well as the set $\{x : f(x) > \alpha\}$ is measurable for every real $\alpha \in \mathbf{R}$. Since the characteristic function of a non-measurable set E denoted by ψ_E *i.e.*, the function which equals 1 at the points of E and 0 otherwise, is non-measurable. Therefore the existence of non-measurable set implies the existence of a non-measurable function.

Theorem 19. *The set* $E \subseteq [ab]$ *and its characteristic function* ψ_E *are both measurable or both non-measurable.*

If ψ_E is measurable then the measurability of E follows from the relation

$$E = \{x : \psi_E > 0\}$$

Conversely if E is a measurable set then the set

$$\{x : \psi_E > \alpha\} = \begin{cases} \phi & \text{if } \alpha \geq 1 \\ E & \text{if } 0 \leq \alpha < 1 \\ [a, b] & \text{if } \alpha < 0 \end{cases}$$

which establishes the measurability of ψ_E.

7. LEBESGUE INTEGRAL

The definition of the integral by Cauchy and Riemann turns out to be inadequate from a more general point of view. First, Riemann's definition has the drawback of applying only rarely; in other words the class of all Riemann integrable functions is quite small. Secondly, the limiting operations often lead to great difficulties. In fact, if $f_1, f_2, ..., f_n$ ... is a sequence of Riemann integrable functions converging pointwise to a function f on $[a, b]$ then it is not in general true that

$$\lim_{n \to \infty} \int_a^b f_n = \int_a^b f$$

To enlarge the class of integrable functions, first we will give the definition of the integral which is similar to that of Riemann. Later on we will define it independently.

Definition 1. Let f be any bounded function on $[a, b]$ and let $P = (A_1, A_2, ... A_n)$ be any partition of $[a, b]$, where $A_1, A_2, ... A_n$ are measurable subsets of $[a, b]$ (also called the components of P) such that

$$\bigcup_{i=1}^{n} A_i = [a, b], \text{ and } m(A_i \cap A_j) = 0, \text{ for } i \neq j$$

Such a partition of $[a, b]$ would be called a *measurable partition* of $[a, b]$. We define

$$U(P, f) = \sum_{i=1}^{n} (\sup_{x \in A_i} f(x))\, m A_i$$

and

$$L(P, f) = \sum_{i=1}^{n} (\inf_{x \in A_i} f(x))\, m A_i$$

as the upper and lower Lebesgue sums of the function f corresponding to the partition P $(A_1, A_2,..., A_n)$ of $[a, b]$. Obviously $U(P, f) \geq L(P, f)$ for every partition P. The infimum of the set of all upper Lebesgue sums is called the *Upper Lebesgue Integral* denoted as:

$$L\overline{\int_a^b} f\, dx = \inf U(P, f) \ \ \forall \text{ partitions } P$$

The supremum of the set of all lower Lebesgue sums is called the *Lower Lebesgue Integral* denoted as:

$$L\underline{\int_a^b} f\, dx = \sup L(P, f) \ \ \forall \text{ partitions } P$$

Definition 2. A bounded function f on $[a, b]$ is said to be *Lebesgue Integrable* if

$$L\underline{\int_a^b} f\, dx = L\overline{\int_a^b} f\, dx$$

In this case we define

$$L\int_a^b f\, dx = L\underline{\int_a^b} f\, dx = L\overline{\int_a^b} f\, dx^*$$

The fact that f is Lebesgue integrable, we express by writing $f \in L[a, b]$.

Lemma. *Let f be a bounded function on $[a, b]$. Then for any two measurable partitions P_1, P_2 of $[a, b]$, we have*

$$U(P_1, f) \geq L(P_2, f), \text{ and } L\underline{\int_a^b} f\, dx \leq L\overline{\int_a^b} f\, dx$$

Let $P_1 = \{A_1, A_2, ..., A_n\}$, $P_2 = \{B_1, B_2, ..., B_m\}$ be any two measurable partitions of $[a, b]$, and P be the measurable partition, called their common refinement whose components are nm subsets $A_i \cap B_j\, (1 = 1, 2, 3, ..., n;\ j = 1, 2, ..., m)$. Now

$$U(P_1, f) \geq U(P, f) \geq L(P, f) \geq L(P_2, f)$$

Taking the infimum over all partitions P_1, we get $\inf\limits_{P_1} U(P_1, f) \geq L(P_2, f)$; whatever partition P_2 may be.

Again taking the supremum over all partitions P_2 we have

$$\inf_{P_1} U(P_1, f) \geq \sup_{P_2} L(P_2, f)$$

i.e.,

$$L\overline{\int_a^b} f\, dx \geq L\underline{\int_a^b} f\, dx$$

Remark: Every upper (lower) Riemann integral is greater than (less than) or equal to every upper (lower) Lebesgue integral.

*For simplicity sometimes we denote the upper and lower Lebesque integrals of f by

$$L\overline{\int_a^b} f \text{ and } L\underline{\int_a^b} f \text{ and the integrals by } L\int_a^b f$$

CHAPTER 20

Since $R\overline{\int_a^b} f\ dx = \inf\limits_Q U(Q, f)$, where Q is any partition of $[a, b]$ into intervals, every interval being a measurable set, so every partition can be regarded as a measurable partition. Thus

$$\inf_Q U(Q, f) \geq \inf_P U(P, f) = L\overline{\int_a^b} f\ dx$$

where P is a measurable partition. Similarly

$$L\underline{\int_a^b} f\ dx \geq R\underline{\int_a^b} f\ dx$$

Combining the two results, we have

$$R\underline{\int_a^b} f\ dx \leq L\underline{\int_a^b} f\ dx \leq L\overline{\int_a^b} f\ dx \leq R\overline{\int_a^b} f\ dx$$

Theorem 20. *Every bounded Riemann integrable function over $[a, b]$ is Lebesgue integrable and the two integrals are equal.*

If f is Riemann integrable, then

$$R\underline{\int_a^b} f\ dx = R\overline{\int_a^b} f\ dx = R\int_a^b f\ dx$$

By the previous remark, we have

$$R\int_a^b f\ dx \leq L\underline{\int_a^b} f\ dx \leq L\overline{\int_a^b} f\ dx \leq R\int_a^b f\ dx$$

$$\Rightarrow \qquad R\int_a^b f\ dx = L\underline{\int_a^b} f\ dx = L\overline{\int_a^b} f\ dx = L\int_a^b f\ dx$$

Thus every Riemann integrable function is Lebesgue integrable and the two integrals are the same. However, the converse of this result may or may not be true. It can be illustrated by the following example.

Let f be a function defined on the interval $[0, 1]$ as follows:

$f(x) = 0$, when x is rational

$ = 1$, when x is irrational

This function is not Riemann integrable (see Example 2, Chapter 9). For Lebesgue integrability, let A_1 be the set of all rational numbers and A_2 be the set of all irrational numbers in $[0, 1]$. The partition $P = \{A_1, A_2\}$ is a measurable partition of $[0, 1]$ and $m A_1 = 0$, $m A_2 = 1$.

$$L(P, f) = \inf_{A_1} f(x) \cdot m A_1 + \inf_{A_2} f(x) \cdot m A_2 = 0 \cdot m A_1 + 1 \cdot m A_2 = 1$$

$$U(P, f) = \sup_{A_1} f(x) \cdot m A_1 + \sup_{A_2} f(x) \cdot m A_2 = 0 \cdot m A_1 + 1 \cdot m A_2 = 1$$

$$\sup_P L(P, f) = 1 = \inf_P U(P, f)$$

$\Rightarrow \qquad f$ is Lebesgue integrable over $[0, 1]$ and the integral of f is 1.

In fact we have the following theorem:

A function f is Riemann integrable on $[a, b]$ if and only if the set of discontinuities of f in $[a, b]$ has measure zero, i.e., if f is continuous a.e.

(The proof is beyond the scope of the present chapter.)

Theorem 21. *A necessary and sufficient condition for a bounded function f to be Lebesgue integrable over $[a, b]$ is that for each given $\in > 0$ there exists a measurable partition P of $[a, b]$ such that $U(P, f) - L(P, f) < \in$.*

[The proof is exactly similar to Theorem 3, Ch. 9.]

Theorem 22. *Every bounded measurable function on $[a, b]$ is Lebesgue integrable on $[a, b]$.*

f being bounded, there exist real numbers m, M such that the range of f is a subset of $[m, M[$. Given $\in > 0$ there exists a finite number of points $y_0, y_1, y_2, ..., y_n$ such that

$$m = y_0 < y_1 < y_2 ... < y_n = M$$

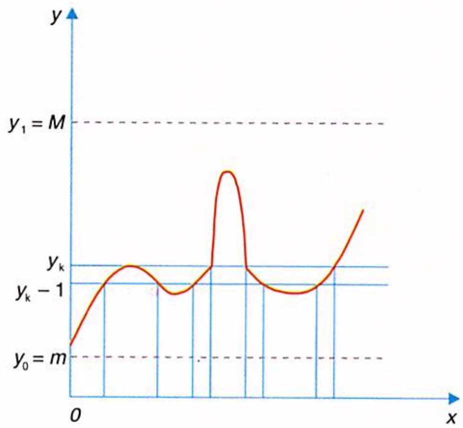


Fig. 1

and

$$y_k - y_{k-1} < \frac{\in}{b-a}$$

For each $k = 1, 2, ..., n$, let $A_k = \{x \in [a, b] : y_{k-1} \le f(x) < y_k\}$. Then each A_k is measurable, since f is measurable.

Thus, $P = \{A_1, A_2, ..., A_n\}$ is a measurable partition of $[a, b]$.

Since $\sup_{x \in A_k} f(x) \le y_k$, and $\inf_{x \in A_k} f(x) \ge y_{k-1}$,

$$U(P, f) = \sum_{k=1}^{n} (\sup_{A_k} f(x))\, mA_k \le \sum_{k=1}^{n} y_k\, mA_k$$

$$L(P, f) = \sum_{k=1}^{n} (\inf_{A_k} f(x))\, mA_k \ge \sum_{k=1}^{n} y_{k-1}\, mA_k$$

Now

$$U(P, f) - L(P, f) \le \sum_{k=1}^{n} (y_k - y_{k-1})\, mA_k < \frac{\in}{b-a} \sum_{k=1}^{n} mA_k < \in.$$

[since A_k's are pairwise disjoint sets and $\bigcup_{k=1}^{n} A_k = [a, b]$.]

Hence f is Lebesgue integrable on $[a, b]$. Thus if f is bounded on $[a, b]$, the measurability of f is a sufficient condition for f to be Lebesgue integrable on $[a, b]$,

It may be noted that measurability is also a necessary condition for a bounded function f to be Lebesgue integrable.

8. PROPERTIES OF LEBESGUE INTEGRAL FOR BOUNDED MEASURABLE FUNCTIONS

The following properties hold for a bounded Lebesgue integrable function f on $[a, b]$:

(*i*) If $a < c < b$, then f is Lebesgue integrable on $[a, c]$ as well as on $[c, b]$ and

$$\int_a^b f = \int_a^c f + \int_c^b f$$

(*ii*) If k *is* any scalar, $k \in \mathbf{R}$, then kf is Lebesgue integrable and

$$\int_a^b kf = k\int_a^b f$$

(*iii*) If $f = f_1 + f_2$, where f_1, f_2 are Lebesgue integrable on $[a, b]$, then f is Lebesgue integrable on $[a, b]$, and

$$\int_a^b f = \int_a^b f_1 + \int_a^b f_2$$

(*iv*) If A_k is a (finite or infinite) sequence of disjoint measurable subsets of $[a, b]$ whose union A has finite measure, then

$$\int_A f = \sum_k \int_{A_k} f$$

Example 10. Let

$$f(x) = \begin{cases} 1, & \text{if } x \text{ is an irrational number in } [-4, 4] \\ -2, & \text{if } x \text{ is rational number in } [-4, 4] \end{cases}$$

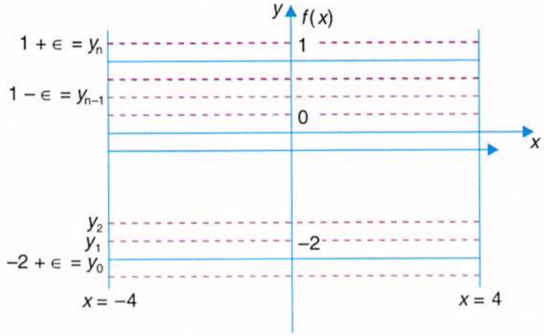


Fig. 2

Evaluate $\int_{-4}^{4} f(x)\,dx$.

As f is bounded by the numbers –2, 1, choose a partition P by subdividing [–2, 1] by points y_n's such that

$$-2 - \in = y_0 < y_1 < y_2 \ldots < y_{n-1} < y_n = 1 + \in, \quad \in > 0$$

and consider the sets A_k's for $k = 1, 2, \ldots, n$ as follows:

$$A_k = \{x : y_{k-1} \le f(x) < y_k\}$$

In the figure the bold lines indicate the graph of f in [–4, 4] and the dotted lines show the points of sub-division taken along y-axis. It is clear from the diagram that the only non-empty sets are

$$A_1 = \{x : y_0 \le f(x) < y_1\}$$

and

$$A_n = \{x : y_{n-1} \le f(x) < y_n\}$$

The set A_1 is the set of rational numbers in [–4, 4] while the set A_n is the set of irrational numbers in [–4, 4]. It follows therefore that $m(A_1) = 0$ and $m(A_n) = 8$, while $m A_k = 0$, for $k = 2, 3, 4, \ldots, (n-1)$.

For the partition $P = \{A_1, A_2, A_3, \ldots, A_n\}$ of measurable subsets of [–4, 4], the upper and lower Lebesgue sums are defined respectively by

$$U(P f) = \sum_{k=1}^{n} \sup_{x \in A_k} f(x)\, m A_k \le \sum_{k=1}^{n} y_k\, m A_k = y_1\, m A_1 + y_n\, m A_n$$

$$= 8 y_n = 8(1 + \in)$$

$$L(P, f) = \sum_{k=1}^{n} \inf_{y \in A_k} f(x)\, m A_k \ge \sum_{k=1}^{n} y_{k-1}\, m A_k = 8 y_{n-1} = 8(1 - \in)$$

Now by varying the partition, y_n is any number greater than one and y_{n-1} is any number less than one.

$$\therefore \qquad \inf_P U(P, f) = \sup_P L(P, f) = 8$$

$$\therefore \qquad \int_{-4}^{4} f(x)\,dx = 8.$$

Example 11. Calculate Lebesgue integral for the function

$$f(x) = \begin{cases} 1 & \text{when } x \text{ is rational} \\ 2 & \text{when } x \text{ is irrational} \end{cases}$$

■ As f is bounded on **R** (the real line) it is bounded on any subset $[a, b]$ of **R**.
By definition the upper and lower Lebesgue sums are given as in the previous example.

$$U(P, f) = \sum_{k=1}^{n} \sup_{A_k} f(x)\, m A_k \le \sum_{k=1}^{n} y_k\, A_k = y_1\, m A_1 + y_n\, m A_n$$

i.e.,

$$U(P, f) \leq (2 + \delta)(b - a) \qquad \because\ mA_1 = 0,\ mA_n = b - a$$

$$L(P, f) = \sum_{k=1}^{n} \inf_{A_k} f(x) \geq \sum_{k=1}^{n} y_{k-1} A_k = y_1\, mA_1 + y_{n-1}\, mA_n$$

i.e.,

$$L(P, f) \geq (2 - \delta)(b - a)$$

where $\qquad A_k = \{x : y_{k-1} \leq f(x) < y_k\},\ k = 1, 2, 3 \ldots n.$

A_1 is the set of rationals in $[a, b]$ which are countable and hence $mA_1 = 0$. A_n is the set of irrationals in $[a, b]$ and $mA_n = b - a$. Taking supremum and infimum over all partitions P, we get

$$2(b - a) \geq \inf_P U(P, f) \geq \sup_P L(P, f) \geq 2(b - a)$$

$$\therefore \qquad \int_a^b f(x)\, dx = 2(b - a)$$

If the interval $[a, b]$ is the whole real line, then

$$\int_R f(x)\, dx = 2m\mathbf{R} = \infty.$$

Example 12. Evaluate

$$\int_0^5 f(x)dx, \text{ if } f(x) = \begin{cases} 0, & 0 \leq x < 1 \\ 1, & \{1 \leq x < 2\} \cup \{3 \leq x < 4\} \\ 2, & \{2 \leq x < 3\} \cup \{4 \leq x < 5\} \end{cases}$$

by using Riemann and Lebesgue definitions of the integral.

- Using Riemann definition of the integral (where the subdivision is taken of the segment [0, 5] by the division points $x_0, x_1, x_2, \ldots, x_n$ on x-axis) the upper and lower Riemann sums tend to the common value

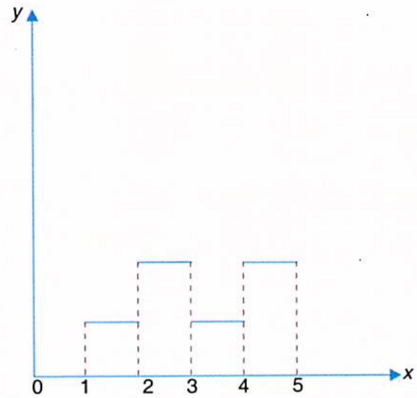


Fig. 3

$$0\,(1-0)+1(2-1)+2(3-2)+1(4-3)+2(5-4)=6$$

(since the function is constant on each of the subintervals.)

$$\therefore \qquad R\int_0^5 f(x)\,dx = 6$$

Evaluating the Lebesgue integral where the subdivision is that of the interval $[0, 2+\delta[, \delta > 0$, we get

$$0\,[1-0]+1\,[(2-1)+(4-3)]+2\,(3-2)+(5-4)]=6$$

$$\therefore \qquad L\int_0^5 f(x)\,dx = 6.$$

Theorem 23. *If f is a bounded and Lebesgue integrable function on* $[a, b]$ *such that* $f(x) = g(x)$ *a.e. on* $[a, b]$, *where g is a bounded function on* $[a, b]$ *then g is Lebesgue integrable and*

$$\int_0^5 f dx = \int_a^b f\,dx$$

f being bounded and Lebesgue integrable, it is measurable. Also $f(x) = g(x)$ a.e. Hence, $g(x)$ is measurable.

Moreover g is bounded, hence g is Lebesgue integrable on $[a, b]$.

To show that $\int_a^b g\,dx = \int_a^b f\,dx$, denote the set A as the set of all points x in $[a, b]$ where $f(x) \neq g(x)$. Clearly $mA = 0$. Now $C(A) = [a, b] - A$ and $f(x) = g(x)$, $x \in C(A)$. Consider the partition $P = \{A, C(A)\}$ of $[a, b]$; then

$$\begin{aligned} U(P, g-f) &= \sup_A (g-f)\cdot mA + \sup_{C(A)} (g-f)\, m\, C(A) \\ &= \sup_A (g-f)\cdot 0 + 0\cdot m\, C(A) = 0 \end{aligned}$$

Similarly $L(P, g-f) = 0$. Thus

$$0 = L(P, g-f) \le \int_a^b (g-f)\,dx \le U(P, g-f) = 0$$

$$\therefore \qquad \int_a^b (g-f)\,dx = 0$$

Now

$$\int_a^b g\,dx = \int_a^b [(g-f)+f]\,dx = \int_a^b (g-f)\,dx + \int_a^b f\,dx = \int_a^b f\,dx$$

$$\therefore \qquad \int_a^b g\,dx = \int_a^b f\,dx\ .$$

Theorem 24. *If f and g are bounded functions and Lebesgue integrable over* $[a, b]$ *and if*

(*i*) $f(x) \ge 0$ *a.e. on* $[a, b]$, *then* $\int_a^b f(x)\,dx \ge 0$.

(*ii*) $f(x) \le g(x)$ *a.e. on* $[a, b]$, *then* $\int_a^b f\,dx \le \int_a^b g\,dx$.

(*i*) We may assume that $f(x) \geq 0$ for all x in $[a, b]$ as this assumption does not alter the value of $\int_a^b f\,dx$ under given conditions.

Obviously $U(P, f) \geq 0$ for any measurable partition P

$\Rightarrow \quad L\bar{\int_a^b} f\,dx \geq 0$, f being Lebesgue integrable

$$L\int_a^b f\,dx = L\bar{\int_a^b} f\,dx \geq 0, \qquad \therefore\ L\int_a^b f\,dx \geq 0$$

(*ii*) Since $f(x) \leq g(x)$ a.e. on $[a, b]$, so $g(x) - f(x) \geq 0$ a.e. on $[a, b]$. By (*i*),

$$\int_a^b (g(x) - f(x))\,dx \geq 0 \Rightarrow \int_a^b g(x)\,dx - \int_a^b f(x)\,dx \geq 0$$

$$\Rightarrow \qquad \int_a^b f\,dx \leq \int_a^b g\,dx.$$

Theorem 25. *If a bounded function f is Lebesgue integrable on* $[a, b]$ *then* $|f|$ *is Lebesgue integrable over* $[a, b]$. *Moreover, if f is Lebesgue integrable then*

$$\left|\int_a^b f\right| \leq \int_a^b |f|$$

f being bounded and Lebesgue integrable on $[a, b]$

Hence f is measurable.

Define $f^+ = \max(f, 0)$, and $f^- = -\min(f, 0)$, then

$$|f| = f^+ + f^- \left(\text{where } \max(f, g) = \frac{1}{2}(f + g) + \frac{1}{2}|f - g|, \text{ and}\right.$$

$$\left.\min(f, g) = \frac{1}{2}\left((f + g) - \frac{1}{2}|f - g|\right)\right)$$

$|f|$ being the sum of two measurable functions is measurable. Also boundedness of f implies boundedness of $|f|$. Therefore $|f|$ is Lebesgue integrable.

Now

$$f \leq |f|, \text{ and } -f \leq |f|$$

$$\therefore \qquad \int_a^b f \leq \int_a^b |f|, \text{ and } -\int_a^b f \leq \int_a^b |f|$$

$$\Rightarrow \qquad \left|\int_a^b f\right| \leq \int_a^b |f|.$$

Example 13. If $f(x)=\dfrac{1}{2}+\sin x,\ 0\le x<2\pi$, find f^+ and f^-.

■
$$f^+=\max(f(x),0)=\begin{cases}\dfrac{1}{2}+\sin x, & 0\le x<\dfrac{7\pi}{6}\\[2mm] 0, & \dfrac{7\pi}{6}\le x\le\dfrac{11\pi}{6}\\[2mm] \dfrac{1}{2}+\sin x, & \dfrac{11\pi}{6}<x<2\pi\end{cases}$$

$$f^-=-\min(f(x),0)=\begin{cases}0, & 0\le x<\dfrac{7\pi}{6}\\[2mm] -\dfrac{1}{2}-\sin x, & \dfrac{7\pi}{6}\le x\le\dfrac{11\pi}{6}\\[2mm] 0, & \dfrac{11\pi}{6}<x<2\pi\end{cases}$$

Thus

$$f=f^+-f^-,\ \text{and}\ |f|=f^++f^-$$

Example 14. If $f(x)=0$, for every x in the Cantor set P_0 and $f(x)=k$ for x in each of the intervals of length $1/3^k$ in G_0, prove that f is Lebesgue integrable on [0, 1] and that $\int_0^1 f=3$.

■ Let P_0 and G_0 be Cantor sets as defined in Example 4, then

$$P_0\cup G_0=[0,1]$$

$f(x)=0,\ \forall x\in P_0$ [the closed intervals which we retain]

$=k\ \forall x$ in each of the intervals of length $1/3^k$ in G_0 [the open intervals which we delete]

From Example 4, if follows that the number of intervals of length $1/3^k$ is 2^{k-1}. Thus

$$\int_0^1 f\,dx=1\cdot\frac{1}{3}\cdot 1+2\cdot\frac{1}{3^2}\cdot 2+3\cdot\frac{1}{3^3}\cdot 2^2+\ldots+k\cdot\frac{1}{3^k}\cdot 2^{k-1}+\ldots$$

$$=\frac{1}{3}\left(1+2\cdot\frac{2}{3}+3\left(\frac{2}{3}\right)^2+\ldots+k\left(\frac{2}{3}\right)^{k-1}+\ldots\right)=3.$$

9. LEBESGUE INTEGRAL OF A BOUNDED FUNCTION OVER A SET OF FINITE MEASURE

Define a function χ on a measurable subset A of $[a,b]$ known as the characteristic function of A as follows:

$$\chi_A(x) = \begin{cases} 1, & x \in A \\ 0, & x \notin A \end{cases}$$

A linear combination $\Psi(x) = \sum_{i=1}^{n} a_i \chi_{A_i}(x), \bigcup_{i=1}^{n} A_i = A$ of the characteristic functions of measurable sets A_i is called a *simple function*. The representation of a simple function is not unique. If $A_i', i = 1, 2, \dots m$ is another collection of measurable sets such that $\bigcup_{i=1}^{m} A_i' = A$, then $\sum_{i=1}^{m} a_i' \chi_{A_i'}(x)$ is also a representation of the given simple function Ψ. More explicitly, given $\Psi = \sum_{i=1}^{n} a_i \chi_{A_i}$, where A_i's are disjoint and $\bigcup_{i=1}^{n} A_i = A$, then $a_1, a_2, \dots a_n$ are distinct non-zero values of Ψ on the sets $A_1, A_2, \dots, A_n$ respectively. The sets A_i are measurable, if Ψ is a measurable function. This representation of Ψ is unique and is called the canonical representation of Ψ. The integral of the measurable function Ψ is defined as

$$\int_A \Psi \, dx = \sum_{i=1}^{n} a_i \, m A_i$$

Definition. For any non-negative measurable function f on a measurable set A, its integral is defined as

$$\int_A f \, dx = \sup_{\phi \le f} \int_A \phi \, dx, \text{ for all simple functions } \phi \text{ on } A$$

Integral of f can also be defined as

$$\int_A f \, dx = \inf_{\Psi \ge f} \int_A \Psi \, dx, \text{ for all simple functions } \Psi \text{ on } A$$

Since f defined on the set A can also be written as $f \cdot \chi_A$,

$$\therefore \qquad \int_A f = \int_A f \cdot \chi_A.$$

Theorem 26. *A necessary and sufficient condition for a bounded function f defined on a measurable set A with finite measure, to be measurable, is that* $\inf_{\Psi \ge f} \int_A \Psi(x) \, dx = \sup_{\phi \le f} \int_A \phi(x) \, dx$, *for all simple functions* ϕ *and* Ψ *on A.*

Necessary Condition. Let f be any bounded measurable function with M as its supremum on the set A. Partitioning the range $]-M, M]$ of the given function f, we define the sets A_i as follows:

$$A_i = \left\{ x : \frac{i-1}{n} M < f(x) \le \frac{iM}{n}, -n \le i \le n \right\}$$

f being measurable the sets A_i are disjoint measurable sets having A as their union, *i.e.*,

$$A = \bigcup_{i=-n}^{n} A_i, \text{ and } mA = \sum_{i=-n}^{n} m A_i$$

The simple functions $\phi_n(x)$ and $\Psi_n(x)$ defined by

$$\phi_n(x) = \frac{M}{n} \sum_{i=-n}^{n} (i-1)\, \chi_{A_i}(x)$$

and

$$\Psi_n(x) = \frac{M}{n} \sum_{i=-n}^{n} i\, \chi_{A_i}(x)$$

satisfy the relation $\phi_n(x) \le f(x) \le \Psi_n(x)$. Thus

$$\inf_{\Psi \ge f} \int_A \Psi(x)\, dx \le \int_A \Psi_n(x)\, dx = \frac{M}{n} \sum_{i=-n}^{n} i\, m A_i$$

and

$$\sup_{\phi \le f} \int_A \phi(x)\, dx \le \int_A \phi_n(x)\, dx = \frac{M}{n} \sum_{i=-n}^{n} (i-1)\, m A_i$$

$$\therefore \qquad 0 \le \inf_{\Psi \ge f} \int_A \Psi(x)\, dx - \sup_{\phi \le f} \int_A \phi(x)\, dx \le \frac{M}{n} \sum_{i=-n}^{n} m A_i = \frac{M}{n} m A$$

Taking n to be an arbitrary large number, we have

$$\inf_{\Psi \ge f} \int_A \Psi(x)\, dx = \sup_{\phi \le f} \int_A \phi(x)\, dx$$

Sufficient Condition: Suppose $\inf_{\phi \ge f} \int_A \Psi(x)\, dx = \sup_{\phi \le f} \int_A \phi(x)\, dx$

$$\therefore \qquad \sup_{\phi \le f} \int_A \phi(x)\, dx - \inf_{\Psi \ge f} \int_A \Psi(x)\, dx = 0$$

$$\Rightarrow \qquad \sup_{\phi \le f} \int_A \phi(x)\, dx + \sup_{\Psi \ge f} \int_A (-\Psi(x))\, dx = 0$$

Thus, for each given n, there are simple functions $\phi_n(x)$ and $\Psi_n(x)$ such that

$$\int_A \Psi_n(x)\, dx - \int_A \phi_n(x)\, dx < \frac{1}{n}$$

Then, the functions $\inf \Psi_n$ and $\sup \phi_n$ are measurable and

$$\sup \phi_n(x) \le f(x) \le \inf \Psi_n(x)$$

Let $F_\alpha = \{x : \sup \phi_n(x) < \inf \Psi_n(x) - 1/\alpha\}$, then

$$\bigcup_{\alpha} F_\alpha = F = \{x : \sup \phi_n(x) < \inf \Psi_n(x)\}$$

But each F_α is contained in the set $\{x : \phi_n(x) < \Psi_n(x) - 1/\alpha\}$ with measure less than α/n. Since n is an arbitrary large number, so $m(F_\alpha) = 0$ and hence $mF = 0$. Thus, $\sup \phi_n(x) = \inf \Psi_n(x)$ except on a set of measure zero and $\sup \phi_n(x) = f = \inf \Psi_n(x)$ a.e. Hence, f is measurable.

Theorem 27. *If f and g are non-negative bounded measurable functions defined on a set $A \subseteq [a, b]$ of finite measure, then*

(*i*) $\int_A \alpha f + \beta g = \alpha \int_A f + \beta \int_A g$, *where α, β are any real numbers.*

(*ii*) *If* $f = g$ *a.e., then* $\int_A f = \int_A g$.

(*iii*) *If* $f \geq 0$ *a.e. then* $\int_A f \geq 0$, *and if* $f \leq g$ *a.e., then* $\int_A f \leq \int_A g$.

(*iv*) *If* $\alpha \leq f(x) \leq \beta$, *then* $\alpha m A \leq \int_A f \leq \beta mA$. *In particular if* $\alpha = \beta$ *then* $\int_A f = \alpha mA$.

(*v*) *If* A_1 *and* A_2 *are disjoint measurable sets of finite measure, then* $\int_{A_1 \cup A_2} f = \int_{A_1} f + \int_{A_2} f$.

(*vi*) *If* $mA = 0$, *and f is measurable, then* $\int_A f = 0$.

(*i*) Let Ψ be any simple function then $\alpha\Psi$ is also a simple function.
If $\alpha > 0$

$$\int_A \alpha f = \inf_{\Psi \geq f} \int_A \alpha\Psi = \alpha \inf_{\Psi \geq f} \int_A \Psi = \alpha \int_A f$$

If $\alpha < 0$,

$$\int_A \alpha f = \inf_{\phi \leq f} \int_A \alpha\phi = \alpha \sup_{\phi \leq f} \int_A \phi = \alpha \inf_{\Psi \geq f} \int_A \Psi = \alpha \int_A f$$

Now if Ψ_1 and Ψ_2 are simple functions greater than or equal to f and g repectively, then $\Psi_1 + \Psi_2$ is also a simple function greater than or equal to $f + g$, and

$$\int_A f + g \leq \int_A (\Psi_1 + \Psi_2) = \int_A \Psi_1 + \int_A \Psi_2.$$

Since infimum on the right hand side is $\int_A f + \int_A g$,

$$\therefore \qquad \int_A f + g \leq \inf_{\Psi_1 \geq f} \int_A \Psi_1 + \inf_{\Psi_2 \geq g} \int_A \Psi_2 = \int_A f + \int_A g$$

On the other hand, if $\phi_1 \leq f$ and $\phi_2 \leq g$, then $\phi_1 + \phi_2$ is a simple function less than or equal to $(f + g)$.

$$\therefore \qquad \int_A (f + g) \geq \int_A \phi_1 + \phi_2 = \int_A \phi_1 + \int_A \phi_2$$

Again supremum on the right hand side is $\int_A f + \int_A g$.

$$\therefore \qquad \int_A f + g \geq \sup_{\phi_1 \leq f} \int_A \phi_1 + \sup_{\phi_2 \leq f} \int_A \phi_2 = \int_A f \int_A g$$

Hence the result follows.

(*ii*) Since $f - g = 0$ a.e., so if Ψ is a simple function greater than or equal to $f - g$, then $\Psi \geq 0$ a.e. and hence

$$\int_A \Psi \geq 0 \Rightarrow \inf_{\Psi \geq (f-g)} \int_A \Psi \geq 0 \Rightarrow \int_A (f - g) \geq 0$$

Similarly taking $\phi \leq (f - g)$, we get

$$\int_A \phi \leq 0 \Rightarrow \sup_{\phi \leq (f-g)} \int_A \phi \leq 0 \Rightarrow \int_A f - g \leq 0$$

$$\therefore \qquad \int_A f - g = 0 \Rightarrow \int_A f = \int_A g$$

(*iii*) $f \geq 0$ a.e. on A. If Ψ is a simple function greater than or equal to f, then

$$\Psi \geq f \geq 0 \text{ a.e.} \Rightarrow \int_A \Psi \geq 0 \Rightarrow \inf_{\Psi \geq f} \int_A \Psi \geq 0 \Rightarrow \int_A f \geq 0$$

$\because$ $$f \leq g \text{ a.e.} \Rightarrow g - f \geq 0 \text{ a.e. on } A$$

Let Ψ be a simple function greater than or equal to $(g - f)$; then

$$\Psi \geq 0 \text{ a.e.} \Rightarrow \int_A \Psi \geq 0 \Rightarrow \inf_{\Psi \geq g-f} \int_A \Psi \geq 0$$

$\Rightarrow$ $$\int_A g - f \geq 0 \Rightarrow \int_A g \geq \int_A f$$

(*iv*) If $\alpha \leq f \leq \beta$, then for Ψ to be a simple function greater than or equal to f, we have

$$\alpha \leq \Psi \Rightarrow \int_A \alpha \leq \int_A \Psi, \text{ or } \alpha\, m A \leq \int_A \Psi$$

Taking infimum on the right,

$$\alpha m A \leq \inf_{\Psi \geq f} \int_A \Psi = \int_A f$$

Similarly,

$$\sup_{\phi \leq f} \int_A \phi \leq \beta\, m\, A$$

Thus,

$$\alpha m A \leq \inf_{\Psi \geq f} \int_A \Psi = \sup_{\phi \leq f} \int_A \phi \leq \beta\, m\, A.$$

(*v*) Since A_1 and A_2 are disjoint measurable sets of finite measure, so their characteristic function $\chi_{A_1 \cup A_2}$ of the union $A_1 \cup A_2$ is measurable and is of finite measure. Moreover $\chi_{A_1 \cup A_2} = \chi_{A_1} + \chi_{A_2}$.

Let Ψ be a simple function greater than or equal to f on A_1 as well as on A_2. Then Ψ will be greater than or equal to f on $A_1 \cup A_2$. Now

$$\int_{A_1 \cup A_2} f\, dx = \inf_{\Psi \geq f} \int_{A_1 \cup A_2} \Psi = \inf_{\Psi \geq f} \int \Psi \chi_{A_1 \cup A_2}$$

$$= \inf_{\Psi \geq f} \int \Psi \chi_{A_1} + \inf_{\Psi \geq f} \int \Psi \chi_{A_2}$$

$$= \int_{A_1} f + \int_{A_2} g$$

Example 15. Show that if f is a non-negative measurable function then $f = 0$ a.e. on a set A iff $\int_A f\, dx = 0$.

■ Let ϕ be any measurable simple function such that $\phi \leq f$. Since $f = 0$ a.e. on A, so $\int_A \phi\, dx \leq 0$. Taking supremum over all those measurable simple functions $\phi \leq f$, we get $\int_A f\, dx \leq 0$. Similarly

$$\inf_{\Psi \geq f} \int_A \Psi\, dx = \int_A f\, dx \geq 0$$

$\therefore$ $$\int_A f\, dx = 0$$

Conversely if $E_n = \left\{x : f(x) > \frac{1}{n}\right\}$, $0 = \int_A f\, dx \geq \int \frac{1}{n} \chi_{E_n}\, dx = \frac{1}{n} m E_n \Rightarrow mE_n = 0$

But $\quad \{x : f(x) > 0\} = \bigcup_{n=1}^{\infty} E_n, \quad \therefore\ m\left(\bigcup_{n=1}^{\infty} E_n\right) = 0$

i.e., $\quad m\{x : f(x) > 0\} = 0, \quad \therefore\ f = 0$ a.e. on A

10. LEBESGUE INTEGRAL FOR UNBOUNDED FUNCTIONS

We now consider Lebesgue integral for unbounded measurable functions.

Let f be a non-negative measurable function on $[a, b]$.

For each $x \in [a, b]$ and $n \in \mathbf{N}$ we define a function $F(x, n)$ as follows:

$$F(x, n) = f(x), \text{ if } 0 \leq f(x) \leq n = n, \text{ if } f(x) > n$$

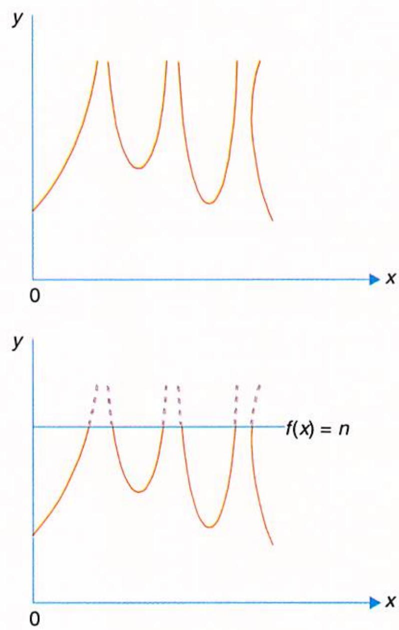


Fig. 4

Thus,

$$F(x, n) = \min(f(x), n)$$

$F(x, n)$ being the minimum of $f(x)$ and n, is bounded and hence measurable which implies that for each $n \in \mathbf{N}$, $F(x, n)$ is Lebesgue integrable.

Now if $\underset{n\to\infty}{\text{Lt}} \int_a^b F(x, n)\, dx$ exists finitely then we say that the unbounded function f is Lebesgue integrable and

$$\int_a^b f = \underset{n\to\infty}{\text{Lt}} \int_a^b F(x, n)\, dx$$

If the limit does not exist finitely then f is not Lebesgue integrable. The function $F(x, n)$ is called *truncated* function.

Example 16. Define

$$f(x) = \frac{1}{x^{2/3}},\ 0 < x \le 1$$
$$0,\ x = 0$$

Show that f is Lebesgue integrable on [0, 1], and $\int_0^1 \frac{1}{x^{2/3}}\, dx = 3$. Find also $F(x, 2)$.

■ Since $\frac{1}{x^{2/3}} \to \infty$, as $x \to 0$, so f is unbounded in [0, 1]. In order to examine its Lebesgue integrability, define

$$\begin{aligned} F(x, n) &= \frac{1}{x^{2/3}},\ \text{if } \frac{1}{n^{3/2}} \le x \le 1 \\ &= n \text{ if } 0 < x < \frac{1}{n^{3/2}} \qquad \ldots(1) \\ &= 0 \text{ if } x = 0 \end{aligned}$$

For $n = 2$

$$\begin{aligned} F(x, 2) &= \frac{1}{x^{2/3}},\ \text{if } \frac{1}{2^{3/2}} \le x \le 1 \\ &= 2 \text{ if } 0 < x < \frac{1}{2^{3/2}} \\ &= 0 \text{ if } x = 0 \end{aligned}$$

Now

$$\begin{aligned} \int_0^1 F(x, n)\, dx &= \int_0^{1/n^{3/2}} F(x, n)\, dx + \int_{1/n^{3/2}}^1 F(x, n)\, dx \\ &= \int_0^{1/n^{3/2}} n\, dx + \int_{1/n^{3/2}}^1 \frac{1}{x^{2/3}}\, dx \\ &= \frac{1}{\sqrt{n}} + 3\left[1 - \left(\frac{1}{n^{3/2}}\right)^{1/3}\right] = 3 - \frac{2}{\sqrt{n}},\ \forall\, n \end{aligned}$$

Thus by the definition of the Lebesgue integral of unbounded functions, we have

$$\int_0^1 f(x)\,dx = \underset{n\to\infty}{\text{Lt}} \int_0^1 F(x, n)\,dx = \underset{n\to\infty}{\text{Lt}} \left(3 - \frac{2}{\sqrt{n}}\right) = 3.$$

Example 17. If

$$f(x) = \frac{1}{x}, \text{ if } 0 < x \le 1$$
$$= 9, \text{ if } x = 0$$

then f is not Lebesgue integrable on $[0, 1]$.

■ Define
$$F(x, n) = \frac{1}{x}, \text{ if } \frac{1}{n} \le x \le 1$$
$$= n, \text{ if } 0 < x < \frac{1}{n}$$
$$= \min(9, n), \text{ if } x = 0$$

$$\int_0^1 F(x, n)\,dx = \int_0^{1/n} n\,dx + \int_{1/n}^1 \frac{1}{x}\,dx = 1 + \log n, \ \forall\, n$$

$$\therefore \qquad \underset{n\to\infty}{\text{Lt}} \int_0^1 F(x, n)\,dx = \underset{n\to\infty}{\text{Lt}} (1 + \log n) = \infty$$

Thus by definition f is not Lebesgue integrable.

11. THE GENERAL INTEGRAL

Let f be any real function defined on $\mathbf{R}$ and A be any measurable subset of $\mathbf{R}$. The functions f^+ and f^- are defined as

$$f^+(x) = \max(f(x)), \text{ and } f^-(x) = \max(-f(x), 0) = -\min(f(x), 0)$$

are such that

(*i*) $f = f^+ - f^-, |f| = f^+ + f^-; f^+, f^- \ge 0$

(*ii*) f is measurable iff f^+ and f^- are measurable.

Definition. Let f be a measurable function defined on a measurable set A and

$\int_A f^+ dx < \infty, \int_A f^- dx < \infty$, then f is Lebesgue integrable on A and its integral is given by

$$\int_A f = \int_A f^+ - \int_A f^-$$

If $\int_A f^+\,dx = +\infty$ and $\int_A f^- dx = +\infty$, then $\int_A f\,dx$ is undefined.

From the definition it follows that $\int_A f$ is finite if both $\int_A f^+ \, dx$ and $\int_A f^- \, dx$ are finite and this will be true iff $\int_A |f| = \int_A f^+ + \int_A f^-$ is finite.

12. SOME FUNDAMENTAL THEOREMS

Theorem 28. *If $A = \bigcup_k A_k$ is the union of a finite or denumerable number of measurable and disjoint subsets $A_1, A_2, A_3, \ldots, A_k, \ldots$ of $[a, b]$, and if $\int_A f \, dx$ exists, then*

$$\int_A f \, dx = \sum_k \int_{A_k} f \, dx$$

We will establish the theorem for non-negative functions and extend the results to arbitrary functions by decomposing the function into non-negative and non-positive parts, *i.e.*,

$$\int_A f \, dx = \int_A f^+ \, dx - \int_A f^- \, dx = \sum_k \left(\int_{A_k} f^+ \, dx - \int_{A_k} f^- \, dx \right)$$

$$= \sum_k \int_{A_k} f \, dx.$$

Hence, we assume that $f \geq 0$.

Let $E_k = A_1 \cup A_2 \ldots \cup A_k$.

For bounded functions $F(x, n) = \min (f(x), n)$, we have

$$\int_A F(x, n) \, dx = \sum_k \int_{A_k} F(x, n) \, dx \leq \sum_k \int_{A_k} f(x) \, dx$$

For every n, when $n \to \infty$, it follows that

$$\int_A f \, dx = \operatorname*{Lt}_{n \to \infty} \int_A F(x, n) \, dx \leq \sum_k \int_{A_k} f \, dx \qquad \ldots(1)$$

Also

$$\int_A f \, dx \geq \int_{E_k = \bigcup_{i=1}^{k} A_i} f \, dx \geq \int_{\bigcup_{i=1}^{k} A_i} F(x, n) \, dx = \sum_{i=1}^{k} \int_{A_i} F(x, n) \, dx$$

Taking the limit first with respect to n and then with respect to k (if $\bigcup_k A_k$ is countable), we get

$$\int_A f \, dx \geq \sum_{i=1}^{k} \int_{A_i} f \, dx \text{ or } \int_A f \, dx \geq \sum_{i=1}^{\infty} \int_{A_i} f \, dx \qquad \ldots(2)$$

according as k is finite or infinite.

From (1) and (2), we have

$$\int_A f \, dx = \sum_k \int_{A_k} f \, dx$$

Remark: If the sequence of sets A_k is finite then the assumption on the existence of the integral $\int_{A_\infty} f(x) \, dx$ is not necessary. The result of the theorem need not hold if we relax the condition of integrability of f on $A = \bigcup_{k=1}^{\infty} A_k$. This can be illustrated by the following example.

Example 18. Let

$$f(x)=\begin{cases} n & \text{if } \dfrac{2n}{4n^2-1}<x\le\dfrac{1}{2n-1} \\ -n & \text{if } \dfrac{1}{2n+1}<x\le\dfrac{2n}{4n^2-1}\end{cases}\qquad n=1,2,3,\ldots$$

For each n we have

$$\int_{1/(2n+1)}^{1/(2n-1)} f(x)\,dx=\int_{1/(2n+1)}^{2n/(4n^2-1)}(-n)\,dx+\int_{2n/(4n^2-1)}^{1/(2n-1)} n\,dx$$

$$=-n\left(\frac{2n}{4n^2-1}-\frac{1}{2n+1}\right)+n\left(\frac{1}{2n-1}-\frac{2n}{4n^2-1}\right)=0$$

and

$$\bigcup_n\left\{\left]\frac{1}{2n+1},\frac{2n}{4n^2-1},\right]\cup\left]\frac{2n}{4n^2-1},\frac{1}{2n-1}\right]\right\}=\,]0,1]$$

Now

$$\int_0^1|f(x)|\,dx=\sum_{n=1}^{\infty}\int_{1/(2n+1)}^{1/(2n-1)}|f(x)|\,dx=\sum_{n=1}^{\infty}\int_{1/(2n+1)}^{1/(2n-1)} n\,dx$$

$$=\sum_{n=1}^{\infty}\frac{2n}{4n^2-1}\ge\sum_{n=1}^{\infty}\frac{1}{4n}$$

But the series $\dfrac{1}{4}\sum_{n=1}^{\infty}\dfrac{1}{n}$ is divergent.

$\therefore\ \int_0^1|f(x)|\,dx$ tends to infinity, and hence $\int_0^1 f(x)\,dx$ does not exist.

Theorem 29. *Let f be a measurable function over the interval $[a, b]$, then f is Lebesgue integrable over $[a, b]$ iff $|f|$ is Lebesgue integrable. Moreover if f is Lebesgue integrable over $[a, b]$, then*

$$\left|\int_a^b f\right|\le\int_a^b|f|$$

Suppose f is measurable and $|f|$ is Lebesgue integrable over $[a, b]$. By definition of f^+ and f^-, we have for $x\in[a,b]$,

$$0\le f^+(x)\le|f(x)|$$
$$0\le f^-(x)\le|f(x)|$$

which implies that f^+ and f^- are Lebesgue integrable over $[a, b]$. Hence f is Lebesgue integrable on $[a, b]$. Again if f is Lebesgue integrable over $[a, b]$ then both f^+ and f^- are Lebesgue integrable.

But $|f|=f^++f^-$. Hence $|f|$ is Lebesgue integrable. Now

$$f \leq |f|, \text{ and } -f \leq |f|$$

$\therefore$ $$\int_a^b f \leq \int_a^b |f| \text{ and } -\int_a^b f \leq \int_a^b |f|$$

$\Rightarrow$ $$\left|\int_a^b f\right| \leq \int_a^b |f|$$

Ex. If f is Lebesgue integrable on $[a, b]$, and $\left|\int_a^b f\, dx\right| \leq \int_a^b |f|\, dx$, when does the equality hold?

$\Big[$*Hint*: $|f| - f \geq 0$ so, $\int |f|\, dx \geq \int |f|\, dx$. Also $|f| + f \geq 0$ and hence $\int_a^b |f|\, dx \geq -\int_a^b f\, dx$.

Hence, $\int_a^b |f|\, dx \geq \left|\int_a^b f\, dx\right|$.

If $\int_a^b f\, dx \geq 0$, then $\int_a^b |f|\, dx = \int_a^b f\, dx$, i.e., $\int_a^b (|f| - f)\, dx = 0 \Rightarrow |f| = f$ a.e.,

i.e., $f \geq 0$ a.e.

If $\int_a^b f\, dx < 0$, then $\int_a^b |f|\, dx = -\int_a^b f\, dx$, *i.e.*, $\int_a^b (|f| + f)\, dx = 0$, and hence $|f| = -f$ a.e., *i.e.* $f \leq 0$ a.e.

Hence, $f \geq 0$ a.e., or $f \leq 0$ a.e. is a necessary condition for the equality of $\left|\int_a^b f\, dx\right|$ *and* $\int_a^b |f|\, dx.\Big]$

Theorem 30. *Let f be a Lebesgue integrable function on $[a, b]$, then given $\in > 0\ \exists$ a $\delta > 0$ such that $\left|\int_A f\right| < \in$ whenever A is a measurable subset of $[a, b]$ with $mA < \delta$.*

We will first prove the result for non-negative functions. Assuming f to be a non-negative function we have

$$\underset{n \to \infty}{\text{Lt}} \int_a^b F(x, n) = \int_a^b f(x)\, dx, \text{ where } F(x, n) = \begin{cases} f(x), & \text{if } 0 \leq f(x) \leq n \\ n, & \text{if } f(x) > n \end{cases}$$

Then given $\in > 0, \exists\, N \in N$ such that

$$\int_a^b f(x)\, dx - \int_a^b F(x, N)\, dx < \frac{\in}{2}$$

$\Big[\because \left\{\int_a^b F(x, n)\, dx\right\}$ is a non-decreasing sequence converging to $\int_a^b f(x)\, dx\Big]$ *i.e.*,

$$\int_a^b (f(x) - F(x, N))\, dx < \frac{\in}{2} \qquad \text{...(1)}$$

Choose any $\delta > 0$, with $\delta < \dfrac{\in}{2N}$. If A is a measurable subset of $[a, b]$ and $mA < \delta$, then we have

$$\int_A F(x, N)\, dx \le \int_A N\, dx = N\, mA < \frac{\in}{2} \qquad ...(2)$$

[using $mA < \in/2N$]

From (1) and (2), we get

$$\int_A f(x)\, dx = \int_A (f(x) - F(x, N))\, dx + \int_A F(x, N)\, dx < \frac{\in}{2} + \frac{\in}{2}$$

For an arbitrary Lebesgue integrable function f over $[a, b]$ we have $f = f^+ - f^-$

By the first part, for a given $\in > 0, \exists$ a $\delta_1 > 0$ such that $\int_A f^+\, dx < \frac{\in}{2}$ whenever $mA < \delta_1$.

Similarly $\exists$ a $\delta_2 > 0$ such that $\int_A f^-\, dx < \frac{\in}{2}$ whenever $mA < \delta_2$.

Thus if $mA < \delta = \min(\delta_1, \delta_2)$, we get

$$\left|\int_A f\right| \le \int_A |f| = \int_A f^+ + \int_A f^- < \frac{\in}{2} + \frac{\in}{2} = \in$$

Hence, $\left|\int_A f\right| < \in$, whenever $mA < \delta$.

13. LEBESGUE THEOREM ON BOUNDED CONVERGENCE

Let $\{f_n\}$ be a sequence of functions measurable on a measurable subset $A \subseteq [a, b]$ such that $\underset{n \to \infty}{\text{Lt}} f_n(x) = f(x)$. Then, if there exists a constant M such that $|f_n(x)| \le M$, for all n and for all x, we have

$$\underset{n \to \infty}{\text{Lt}} \int_A f_n(x)\, dx = \int_A f(x)\, dx$$

Since $\underset{n \to \infty}{\text{Lt}} f_n(x) = f(x)$, and $|f_n(x)| \le M \Rightarrow |f(x)| \le M$, the function f is bounded and measurable (being the limit of bounded and measurable functions), hence Lebesgue integrable. We shall show that

$$\underset{n \to \infty}{\text{Lt}} \int_A |f_n(x) - f(x)|\, dx = 0$$

For a given $\in > 0$, we define a partition A into disjoint measurable sets A_k's as follows:

$$A_k = \left\{x : |f_{k-1} - f| \ge \in, |f_n - f| < \in, \forall\, n \ge k\right\},\ k = 1, 2, 3, ...$$

In particular,

$$A_1 = \left\{x : |f_n - f| < \in, n = 1, 2, ...\right\}$$
$$A_2 = \left\{x : |f_1 - f| \ge \in, |f_n - f| < \in, n = 2, 3, 4, ...\right\}$$

Clearly

$$A = \bigcup_{k=1}^{\infty} A_k = \left(\bigcup_{k=1}^{\infty} A_k \right) \cup \left(\bigcup_{k=n+1}^{\infty} A_k \right) = P_n \cup Q_n \quad \text{(say)}$$

and

$$m\,A = m\,(P_n \cup Q_n) = m\,P_n + m\,Q_n$$

Now

$$\int_A |f_n - f|\,dx = \int_{P_n} |f_n - f|\,d\,x + \int_{Q_n} |f_n - f|\,dx \qquad \text{...(1)}$$

For each n, we have

$$|f_n - f| < \varepsilon \text{ on } P_n, \text{ and } |f_n - f| \le |f_n| + |f| \le 2M \text{ on } Q_n \qquad \text{...(2)}$$

Thus

$$\int_A |f_n - f|\,d\,x < \varepsilon\,m\,P_n + 2M\,m\,Q_n$$

As $n \to \infty$, $\underset{n \to \infty}{\text{Lt}}\, m\,P_n = m\,A$ and, $\underset{n \to \infty}{\text{Lt}}\, m\,Q_n = 0$.

Thus $\quad \int_A |f_n - f|\,dx < \varepsilon\,m\,A$

$\therefore \quad \left|\int_A (f_n - f)\right| \le \int_A |f_n - f| < \varepsilon\,m\,A$, ε being arbitrary

$\therefore \quad \underset{n \to \infty}{Lt} \int_A f_n(x)\,dx = \int_A f(x)\,dx$.

Example 19. Verify Bounded Convergence Theorem for the sequence of functions

$$f_n(x) = \frac{1}{(1 + x/n)^n},\ 0 \le x \le 1,\ n \in \mathbf{N}.$$

$$|f_n(x)| = \left|\frac{1}{(1 + x/n)^n}\right| \le 1,\ \forall\, n \text{ and } \forall\, x.$$

- Each f_n being bounded and measurable, the limit function

$$\underset{n \to \infty}{\text{Lt}}\, f_n(x) = \underset{n \to \infty}{\text{Lt}}\, \frac{1}{(1 + x/n)^n} = \frac{1}{e^x}$$

is also bounded and measurable.

Now, $$\int_0^1 \frac{d\,x}{(1 + x/n)^n} = \left.\frac{n\,(1 + x/n)^{-n+1}}{(-n+1)}\right|_0^1 = \frac{n}{n-1}\left(1 - \frac{(1 + 1/n)}{(1 + 1/n)^n}\right)$$

$\therefore$ $$\underset{n \to \infty}{\text{Lt}} \int_0^1 \frac{d\,x}{(1 + x/n)^n} = \underset{n \to \infty}{\text{Lt}}\, \frac{n}{n-1}\left(1 - \frac{1 + 1/n}{(1 + 1/n)^n}\right)$$

$$0 = \left(1 - \frac{1}{e}\right) = \frac{e-1}{e}$$

Also

$$\int_0^1 \underset{n \to \infty}{\text{Lt}} \frac{1}{(1 + x/n)^n}\, dx = \int_0^1 \frac{1}{e^x}\, dx = \left(1 - \frac{1}{e}\right) = \frac{e-1}{e}$$

Hence, Bounded Convergence Theorem is applicable.

Theorem 31. ***Monotone convergence theorem.*** *Let A be any measurable subset of $[a, b]$ with finite measure. Let $\{f_n\}$ be a sequence of measurable functions such that for $x \in A$*

$$0 \le f_1(x) \le f_2(x) \le \ldots \le \ldots f_n(x) \le \ldots$$

If $\underset{n \to \infty}{Lt} f_n(x) = f(x)$, then $\underset{n \to \infty}{Lt} \int_A f_n = \int_A f$.

For a given x, the sequence $\{f_n(x)\}$ is monotonically increasing and $\underset{n \to \infty}{\text{Lt}} f_n(x) = f(x)$. Assume that $\underset{n \to \infty}{\text{Lt}} \int_A f_n = \alpha$, for some α.

Since $\int_A f_n \le \int_A f$, so taking the limit as $n \to \infty$, we get

$$\alpha \le \int_A f \qquad \ldots(1)$$

Let β be any real number lying between 0 and 1, *i.e.*, $0 < \beta < 1$ and ϕ be any simple measurable function such that $0 \le \phi \le f$.

Define the sets $A_n = \{x : f_n(x) \ge \beta\phi(x)\}$, $n = 1, 2, 3 \ldots$

By definition of $\{f_n\}$, we have

$$A_i \subseteq A_{i+1},\ i = 1, 2, \ldots, n, \ldots \text{ and } A = \bigcup_{i=1}^{\infty} A_i$$

Also for any n,

$$\int_A f_n \ge \int_{A_n} f_n \ge \beta \int_{A_n} \phi$$

Taking the limit as n tends to infinity, we get

$$\alpha \ge \beta \int_A \phi$$

Letting $\beta \to 1$, we get $\alpha \ge \int_A \phi$. Since ϕ is arbitrary, so the result holds for all simple functions $\phi \le f$.

$$\therefore \qquad \alpha \ge \sup_{\phi \le f} \int_A \phi = \int_A f \qquad \ldots(2)$$

The equality follows from (1) and (2).

14. INTEGRABILITY AND MEASURABILITY

The next theorem establishes the connection between integrability and measurability and, indeed, provides a major justification for introducing the concept of measurability of functions. We define a function f to be dominated in A by a function g if g is integrable and $|f| \leq g$ throughout A.

Theorem 32. ***Classical Lebesgue dominated convergence theorem.*** *Let $\{f_n\}$ be a sequence of measurable functions on $[a, b]$ such that*

$$\underset{n \to \infty}{\text{Lt}}\ f_n(x) = f(x) \text{ a.e. on } [a, b]$$

If there exists a Lebesgue integrable function g on $[a, b]$ such that for each $n \in \mathbf{N}$

$$|f_n(x)| \leq g(x) \text{ a.e. on } [a, b]$$

Then f is Lebesgue integrable and

$$\underset{n \to \infty}{Lt} \int_a^b f_n = \int_a^b f$$

f being the limit of a sequence of measurable functions is also measurable. For $n \in I$

$$|f_n(x)| \leq g(x) \text{ a.e. on } [a, b] \qquad \ldots(1)$$

$$\Rightarrow \qquad \underset{n \to \infty}{\text{Lt}} |f_n(x)| \leq g(x) \text{ a.e.} \Rightarrow \left| \underset{n \to \infty}{\text{Lt}}\ f_n(x) \right| \leq g(x) \text{ a.e.}$$

$$\Rightarrow \qquad |f(x)| \leq g(x) \text{ a.e. on } [a, b] \qquad \ldots(2)$$

Since g is Lebesgue integrable on $[a, b]$, so $|f|$ is Lebesgue integrable on $[a, b]$, which in turn implies that f is Lebesgue integrable on $[a, b]$.

Corresponding to a given $\in > 0$ and $N \in \mathbf{N}$, define

$$A_N = \left\{ x : |f_n - f(x)| < \frac{\in}{2(b-a)},\ n \geq N,\ x \in [a, b] \right\}$$

Here

$$A_N \subset A_{N+1} \quad \forall\ N \in \mathbf{N}$$

Since $\underset{n \to \infty}{\text{Lt}}\ f_n(x) = f(x)$ a.e., so every point of $[a, b]$ either lies in the union $\overset{\infty}{\underset{N=1}{\cup}} A_N$ or in a subset A of the interval $[a, b]$ whose measure is zero, *i.e.*

$$[a, b] = \left(\overset{\infty}{\underset{N=1}{\cup}} A_N \right) \cup A, \text{ where } mA = 0$$

$$\therefore \qquad m\left(\overset{\infty}{\underset{N=1}{\cup}} A_N \right) = b - a$$

Also

$$m\left(\overset{\infty}{\underset{N=1}{\cup}} A_N \right) = \underset{N \to \infty}{\text{Lt}}\ m\left(\overset{N}{\underset{n=1}{\cup}} A_n \right), \text{ i.e., } m\left(\overset{\infty}{\underset{N=1}{\cup}} A_N \right) = \underset{N \to \infty}{\text{Lt}}\ m A_N$$

$$\therefore \qquad b - a = \operatorname*{Lt}_{N \to \infty} m A_N \qquad \text{...(3)}$$

It is given that g is Lebesgue integrable; therefore for every $\in > 0$ there exists a $\delta > 0$ such that $\int_A g < \frac{\in}{4}$, for $mA < \delta$

From (i) and (ii), it follows that

$$\int_A |f_n| \le \int_A g < \frac{\in}{4}, \quad \text{for } mA < \delta \qquad \text{...(4)}$$

and

$$\int_A |f| \le \int_A g < \frac{\in}{4}, \quad \text{for } mA < \delta \qquad \text{...(5)}$$

From (iii) there exists a positive integer M, such that

$$b - a - m A_N < \delta, \ \forall \ N \ge M \qquad \text{...(6)}$$

i.e.,

$$m C(A_N) < \delta, \ \forall \ N \ge M$$

Now

$$\int_a^b |f_n - f| = \int_{A_M} |f_n - f| + \int_{C(A_M)} |f_n - f| \qquad \text{(by def. of } A_M)$$

$$< \frac{\in}{2} \frac{m(A_M)}{2(b-a)} + \int_{C(A_M)} |f_n| + \int_{C(A_M)} |f|, \ \forall \ n \ge M$$

or

$$\int_a^b |f_n - f| < \frac{\in m A_M}{2(b-a)} + \frac{\in}{4} + \frac{\in}{4} = \in \ \forall \ n \ge M \qquad \text{(using (4) and (5))}$$

$$\therefore \qquad \left| \int_a^b f_n - f \right| \le \int_a^b |f_n - f| < \in \ \forall \ n \ge M$$

or

$$\left| \int_a^b f_n - \int_a^b f \right| < \in \ \forall \ n \ge M$$

$$\Rightarrow \qquad \operatorname*{Lt}_{n \to \infty} \int_a^b f_n = \int_a^b f$$

Remark: The theorem still holds if the interval $[a, b]$ is replaced by any arbitrary measurable subset $A \subseteq [a, b]$ not necessarily the interval. The dominated convergence theorem will not hold if there does not exist an integrable function g such that $|f_n| \le g$, $n = 1, 2, 3, \ldots$ A sequence of functions $\{f_n(x)\}$ although convergent everywhere, can have a nonintegrable limit function. Even if the limit function is integrable $\operatorname*{Lt}_{n \to \infty} \int_a^b f_n \ dx = \int_a^b f \ dx$ may be false, which can be seen by the following example.

Example 20. Define the functions $f_n(x)$ on [0, 1] as follows:

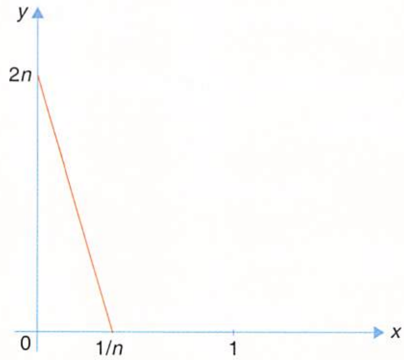


Fig. 5

$$f_n(x) = -2xn^2 + 2n \text{ for } 0 \le x \le \frac{1}{n}$$
$$= 0 \text{ for } \frac{1}{n} < x \le 1$$

$$\int_0^1 f_n(x)\,dx = \int_0^{1/n} \left(-2xn^2 + 2n\right) dx + \int_{1/n}^1 0\,dx$$
$$= 1, \quad n = 1, 2, 3, \ldots$$

$$\therefore \qquad \underset{n \to \infty}{\mathrm{Lt}} \int_0^1 f_n(x)\,dx = 1$$

Also $f_n(x) \to 0$ for every x. Hence $\int_0^1 f(x)\,dx = 0$.

Ex. For $n \ge 2$ define $f_n(x)$ on [0, 1] as

$$f_n(x) = \begin{cases} n^2 x & \text{for } 0 \le x \le 1/n \\ -n^2(x - 2/n) & \text{for } 1/n \le x \le 2/n \\ 0 & \text{for } 2/n \le x \le 1 \end{cases}$$

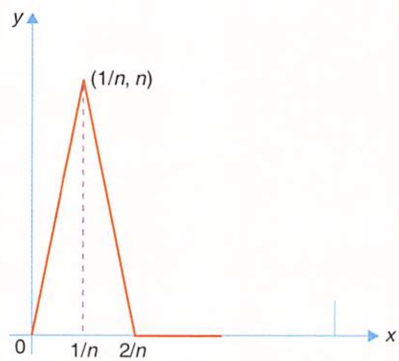


Fig. 6

Show that $f_n \to f$ pointwise on [0, 1] where f is the zero function and

$$\underset{n \to \infty}{\text{Lt}} \int_0^1 f_n(x)\, dx = 1 \neq 0 = \int_0^1 f(x)\, dx.$$

Fatou's Lemma. *Let* $\{f_n\}$ *be a sequence of non-negative measurable functions on* $[a, b]$. *Let* $\lim\limits_{n \to \infty} f_n(x) = f(x)$ a.e. *on* $[a, b]$. *Then*

$$\liminf_{n \to \infty} \int_a^b f_n \geq \int_a^b f$$

if f is Lebesgue integrable on $[a, b]$; *otherwise*

$$\liminf_{n \to \infty} \int_a^b f_n = \infty$$

For any $m \in \mathbf{N}$, we have

$$\begin{aligned} F_n(x, m) &= f_n(x), \text{ if } 0 \leq f_n(x) \leq m, x \in [a, b] \\ &= m, \text{ if } f_n(x) > m \end{aligned}$$

Thus $F_n(x, m) = \min(f_n(x), m)$. Each $F_n(x, m)$ is bounded by m. Also

$$\begin{aligned} \underset{n \to \infty}{\text{Lt}} F_n(x, m) &= \underset{n \to \infty}{\text{Lt}} \min(f_n(x), m) = \min\left(\underset{n \to \infty}{\text{Lt}} f_n(x), m\right) \\ &= \min(f(x), m) \text{ a.e. on } [a, b] \\ &= F(x, m) \text{ a.e. on } [a, b]. \end{aligned}$$

For each m, $F_n(x, m)$ being bounded and measurable, is Lebsegue integrable on $[a, b]$. Thus by the Dominated Convergence Theorem (for $g(x) = m$) its limit viz. $F(x, m)$ is also Lebesgue Integrable. Moreover,

$$\underset{n \to \infty}{\text{Lt}} \int_a^b F_n(x, m)\, dx = \int_a^b \underset{n \to \infty}{\text{Lt}} F_n(x, m)\, dx = \int_a^b F(x, m)\, dx$$

$F_n(x, m)$ is $\min(f_n(x), m)$. Hence $F_n(x, m) \leq f_n(x)$ for all n

$$\begin{aligned} \int_a^b F(x, m)\, dx &= \underset{n \to \infty}{\text{Lt}} \int_a^b F_n(x, m)\, dx \\ &= \underset{n \to \infty}{\text{Lt}} \inf \int_a^b F_n(x, m)\, dx \leq \underset{n \to \infty}{\text{Lt}} \inf \int_a^b f_n \end{aligned}$$

$\therefore$ $$\underset{n \to \infty}{\text{Lt}} \inf \int_a^b f_n \geq \int_a^b F(x, m)\, dx \text{ for each } m \in I$$

$F(x, m)$ tends to $f(x)$ when $m \to \infty$

$\therefore$ $$\liminf_{n \to \infty} \int_a^b f_n \geq \int_a^b f$$

An Important Result for Integrable Functions

Every integrable function f defined on $[a, b]$ can always be *approximated by some continuous function* g for which the following condition holds:

$$\int_a^b |f(x) - g(x)|\, dx < \varepsilon \text{ for any arbitrary positive number } \varepsilon$$

This can be illustrated by the following example.

Define

$$f(x) = \begin{cases} 1, & x \in [a_1, b_1] \subset [a, b] \\ 0, & x \notin [a_1, b_1] \end{cases}$$

Show that for a given $\varepsilon > 0$ there exists a continuous function g defined on $[a, b]$ such that

$$\int_a^b |f(x) - g(x)|\, dx - \varepsilon .$$

Let the graph of g be represented by $PLABMQ$ in the figure; *i.e.*,

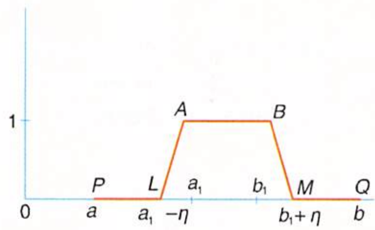


Fig. 7

$$g(x) = \begin{bmatrix} 0, & a \le x \le a_1 - \eta \\ 1 + (x - a_1)/\eta, & a_1 - \eta \le x \le a_1 \\ 1, & a_1 \le x \le b_1 \\ 1 - (x - b_1)/\eta, & b_1 \le x \le b_1 + \eta \\ 0, & b_1 + \eta \le x \le b \end{bmatrix}$$

where η is a +ve small number.

Clearly g is continuous on $[a, b]$. Also

$$\int_a^b |f(x) - g(x)|\, dx = \int_{a_1-\eta}^{a_1} \left(1 + \frac{x - a_1}{\eta}\right) dx + \int_{b_1}^{b_1+\eta} \left(1 - \frac{x - b_1}{\eta}\right) dx = 2\eta < \varepsilon .$$

Example 21. For $n \in \mathbf{N}$ let

$$f_n(x) = 2n, \text{ for } x \in \left]\frac{1}{2n}, \frac{1}{n}\right[$$

CHAPTER 20

$$= 0, \text{ for } \in \left]0, \frac{1}{2n}\right[\cup \left]\frac{1}{n}, 1\right[$$

Calculate $\int_0^1 \lim_{n \to \infty} f_n(x)\, dx$, and $\lim_{n \to \infty} \int_0^1 f_n(x)\, dx$

Show also that Fatou's Lemma holds but that Lebesgue dominated convergence theorem does not.

■ By definition, $\lim_{n \to \infty} f_n(x) = 0$ a.e. on $[0, 1]$

$\therefore$ $$\int_0^1 \lim_{n \to \infty} f_n(x)\, dx = 0$$

Also

$$\int_0^1 f_n(x)\, dx = \int_0^{1/2n} 0 \cdot dx + \int_{1/2^n}^{1/n} 2n\, dx + \int_{1/n}^1 0\, dx = 1$$

$\therefore$ $$\lim_{n \to \infty} \int_0^1 f_n(x) = 1 \Rightarrow \lim_{n \to \infty} \inf \int_0^1 f_n(x)\, dx = 1 > \int_0^1 f\, dx = 0$$

Hence strict inequality holds in Fatou's Lemma. In order that the Lebesgue Dominated Convergence theorem is applicable, the functions $f_n(x)$ should be bounded by some Lebesgue integrable function $g(x)$. But the functions $f_n(x)$ as defined above are unbounded.

Example 22. Put

$$g(x) = \begin{cases} 0, & 0 \le x \le \frac{1}{2} \\ 1, & \frac{1}{2} < x \le 1 \end{cases}$$

$$f_{2k}(x) = g(x), \quad 0 \le x \le 1$$

$$f_{2k+1}(x) = g(1 - x), \quad 0 \le x \le 1$$

Then show that

$$\lim_{n \to \infty} \inf \int_0^1 f_n(x)\, dx > \int_0^1 \lim_{n \to \infty} \inf f_n(x)\, dx$$

■ Since $f_n(x) = g(x)$, $0 \le x \le 1$, if n is even,

$\therefore$ $$f_{2n}(x) = 0, \text{ if } 0 \le x \le \frac{1}{2}$$

$$= 1, \text{ if } \frac{1}{2} < x \le 1$$

$$f_{2n-1}(x) = g(1 - x), \quad 0 \le x \le 1$$

i.e.,

$$f_{2n-1}(x) = 1, \quad 0 \le x < \frac{1}{2}$$

$$= 0, \ \frac{1}{2} \le x \le 1$$

$$\lim_{n \to \infty} \inf f_n(x) = \sup \left\{ \inf \left(f_n, f_{n+1}, f_{n+2} \dots \right) = 0, \ 0 \le x \le 1 \right.$$

$\therefore$ $$\int_0^1 \lim_{n \to \infty} \inf f_n(x)\, dx = 0 \qquad \dots(1)$$

Now

$$\int_0^1 f_{2n}(x) = \int_0^{1/2} 0\, dx + \int_{1/2}^1 1 \cdot dx = \frac{1}{2}$$

and

$$\int_0^1 f_{2n-1}(x)\, dx = \int_0^{1/2} 1 \cdot dx + \int_{1/2}^1 0\, dx = \frac{1}{2}$$

$\therefore$ $$\lim_{n \to \infty} \inf \int_0^1 f_n(x)\, dx = \frac{1}{2} \qquad \dots(2)$$

From (1) and (2),

$$\lim_{n \to \infty} \inf \int_0^1 f_n(x)\, dx > \int_0^1 \lim \inf f_n(x)\, dx$$

Example 23. Let $f_n(x) = \begin{cases} 1/n, & |x| \le n \\ 0, & |x| > n \end{cases}$

Then $f_n(x) \to 0$ uniformly on $\mathbf{R}$ but $\int_{-\infty}^{\infty} f_n\, dx = 2, \ n = 1, 2, 3$

$$\underset{n \to \infty}{\text{Lt}} f_n(x) = \underset{n \to \infty}{\text{Lt}} \frac{1}{n} = 0, \ \text{when } |x| \le n$$

$$= 0, \ \text{when } |x| > n$$

$\therefore \underset{n \to \infty}{\text{Lt}} f_n(x) = 0$, uniformly on the whole real line

Now

$$|f_{2m}(x) - f_m(x)| = \left| \frac{1}{2m} - \frac{1}{m} \right| = \frac{1}{2m} < \epsilon, \ \text{whenever } m > \frac{1}{2\epsilon}$$

Now

$$\int_{-\infty}^{\infty} f_n(x)\, dx = \int_{-\infty}^{-n} 0\, dx + \int_{-n}^{n} \frac{1}{n}\, dx + \int_n^{\infty} 0\, dx = 2$$

This implies that uniform convergence of $\{f_n(x)\}$ is not enough for $\underset{n \to \infty}{\text{Lt}} \int f_n = \int \underset{n \to \infty}{\text{Lt}} f_n$. This equality in Lebesgue integration, in general, is only due to dominated convergence of the sequence

$\{f_n(x)\}$. However, on the set of finite measure, uniformly convergent sequences of bounded functions are boundedly convergent.

Example 24. The function f is defined on the interval [0, 1] by

$$f(x) = \begin{cases} 0, & x \text{ is rational} \\ [1/x], & x \text{ is irrational} \end{cases}$$

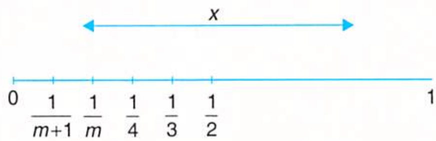
where $[1/x]$ is the integral part of $1/x \leq 1/x$ for x to be irrational. Show that

$$\int_0^1 f\, dx = \infty$$

■ Define a function $g(x) = [1/x]$, for $x \in]0, 1[$. Clearly $f \leq g$ on]0, 1[. Moreover $f = g$ a.e. Since $f < g$ at the rational points of]0, 1[which are of measure zero, thus $f = g$ a.e.

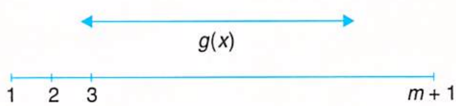
$\therefore$ $$\int_0^1 (g - f)\, dx = 0 \qquad ...(1)$$

Consider $$\int_0^1 g\, dx > \int_{1/(m+1)}^1 g\, dx = \sum_{n=1}^{m} \frac{1}{n+1}$$



The series $\sum_{n=1}^{\infty} \frac{1}{n+1}$ diverges to ∞.

$\therefore$ $$\int_0^1 g\, dx = \infty \qquad ...(2)$$



But

$$\int_0^1 g\, dx = \int_0^1 [(g - f) + f]\, dx = \int_0^1 (g - f)\, dx + \int_0^1 f\, dx$$

$$= \int_0^1 f\, dx \quad [\text{using } (i)]$$

$\therefore$ $$\int_0^1 f\, dx = \infty$$

Example 25. The function f is defined on [0, 1] by

$f(x) = 0$, when x is rational

$= n$, when x is irrational

where n denotes the number of zeros immediately after the decimal point in the representation of x on the decimal scale. Show that f is measurable and find $\int f\ dx$.

■ Let $g(x) = n$, if $\frac{1}{10^{n+1}} \le x < \frac{1}{10^n}$, $n = 0, 1, 2 \ldots$ and $x \in]0, 1]$.

Clearly $f \le g$ over $]0, 1]$. Moreover $f = g$ a.e.

Now the set $\{x : g(x) = n\}$ is measurable (since $g(x)$ is constant on each of the intervals $\left[\frac{1}{10^{n+1}}, \frac{1}{10^n}\right[$, $n = 1, 2, 3$ and each interval is measurable).

Hence the function f is measurable and

$$\int_0^1 f\ dx = \int_0^1 g\ dx$$

But

$$\int_0^1 g\ dx = \sum_{n=0}^{\infty} n\left(\frac{1}{10^n} - \frac{1}{10^{n+1}}\right) = \sum_{n=1}^{\infty} \frac{9n}{10^{n+1}} = \frac{1}{9}$$

$\therefore$
$$\int_0^1 f\ dx = \frac{1}{9}.$$

Example 26. If f is integrable, show that f is finite-value a.e.

■ Let A be any measurable subset of **R**.

As $\int_A f$ is finite, so $\int_A |f|$ is also finite.

Let if possible $|f| = \infty$ on a set $E \subseteq A$ with $mE > 0$, then

$$\int_A |f| > \int_E |f| > n\, m E, \forall\, n \qquad [\because\ |f| > \forall\, n]$$

$\Rightarrow$
$$\int_A |f| \to \infty \text{ as } n \to \infty$$

which is a contradiction to the fact that $\int_A f$ is finite. Thus $mE = 0$. Hence f is finite valued a.e.

Example 27. If f is measurable and g is integrable on $[a, b]$ and M and M' are real numbers such that $M \le f \le M'$ a.e., then $\exists$ a real number ξ. $M \le \xi \le M'$ such that

$$\int_a^b f\,|g|\, dx = \xi \int_a^b |g|\, dx$$

■ Since $|f g| \le |M'|\,|g| \le (|M| + |M'|)\,|g|$ a.e. and g is integrable, so $(|M| + |M'|)\,|g|$ is integrable

$\Rightarrow$ fg is integrable. Also,

$$M|g| \le f|g| \le M'|g| \text{ a.e.}$$

$$\therefore \qquad M \int_a^b |g|\, dx \le \int_a^b f\,|g|\, dx \le M' \int_a^b |g|\, dx$$

If $\int_a^b |g|\, dx = 0$ then $g = 0$ a.e. and the result is trivial. If $\int_a^b |g|\, dx \neq 0$, by taking ξ to be a real number

$$\xi = \frac{\int_a^b f\,|g|\, dx}{\int_a^b |g|\, dx}.$$

we get the required result.

15. LEBESGUE INTEGRAL ON UNBOUNDED SETS OR INTEGRALS

Up till now we have considered Lebesgue integral on bounded sets. In order to extend the definition to unbounded sets such as intervals $]a, \infty]$, $]-\infty, b$ [or] $-\infty, \infty[$, we adopt a suitable limiting case of bounded sets.

Let f be non-negative and Lebesgue integrable on $[a, b]$ for all finite values of b. Then we define Lebesgue integral of f on an unbounded set $]a, \infty[$ as

$$\int_a^\infty f\, dx = \underset{b \to \infty}{\mathrm{Lt}} \int_a^b f\, dx$$

provided this limit exists. If f has an arbitrary sign, then we define

$$\int_a^\infty f\, dx = \int_a^\infty f^+\, dx - \int_a^\infty f^-\, dx$$

where f^+, f^- are non-negative and we say that f is integrable on $]a, \infty[$ if each of the integrals on the right exist in accordance with the definition. Similarly the integrals on the unbounded intervals $]-\infty, b[$ and $]-\infty, \infty[$ can be defined. If A be any unbounded set not necessarily an infinite interval then we can define

$$\int_A f\, dx = \underset{\substack{a \to -\infty \\ b \to +\infty}}{\mathrm{Lt}} \int_{A \cap\,]a,\, b[} f\, dx$$

Most of the theorems which we have established involving Lebesgue integrals on bounded sets also hold for unbounded sets such as Lebesgue dominated convergence theorem, Fatou's Lemma, theorem regarding integrability of f and $|f|$ viz. A function f is integrable on A if and only if $|f|$ is integrable on A regardless of whether A is bounded or unbounded and $\left|\int_A f\right| \le \int_A |f|$.

16. COMPARISON WITH RIEMANN INTEGRAL FOR UNBOUNDED SETS

The Riemann integral of f on unbounded set A can exist eventhough the Riemann integral of $|f|$ does not exist on A. For example, $R\int_0^\infty \frac{\sin x}{x}\, dx = \underset{b \to \infty}{\mathrm{Lt}}\ R\int_0^b \frac{\sin x}{x}\, dx$ exists as an improper Riemann integral

whereas the integral $R\int_0^{\infty}\left|\frac{\sin x}{x}\right|dx$ does not exist. On the contrary the Lebesgue integral of $L\int_0^{\infty}\frac{\sin x}{x}\,dx$ does not exist, because $\int_0^{\infty}\left|\frac{\sin x}{x}\right|dx$ does not exist. It shows that there exist improper Riemann integrals which are not integrable in Lebesgue sense. This indicates *that nothing can be said about the equality of the two integrals when A is unbounded.* Riemann integral may exist when the Lebesgue integral does not exist. Moreover, if $|f|$ is Riemann integrable on A, then f is both Riemann and Lebesgue integrable on A and two integrals are equal. Thus, while talking of Lebesgue integral on unbounded sets, one gets the feeling that it is less general than the improper Riemann integral, yet the great significance of this cannot be ignored due to the most important theorem viz. Lebesgue dominated convergence theorem, which is valid only for it.

EXERCISE

1. Find the lengths of the following sets:
   (*i*) $\{x : -3 < x < 4\} \cup \{x : 1 \le x \le 6\}$
   (*ii*) $\{x : 3 \le x \le 4\} \cup \{-5 \le x \le -2\}$
   (*iii*) $\left\{x : \frac{1}{3^k} \le x \le \frac{1}{3^{k-1}}\right\}$.
2. Prove that the set of all irrational numbers in [0, 1] is measurable and find its measure. Is every subset of irrational numbers measurable?
3. Is every subset of a measurable set measurable?
   [*Hint*: Define a non-measurable subset of [0, 1].]
4. If A_1, A_2, A_3 are any measurable sets, prove that
$$m(A_1 \cup A_2 \cup A_3) = (mA_1 + mA_2 + mA_3) - m(A_1 \cap A_2) - m(A_1 \cap A_3) - m(A_2 \cap A_3) + m(A_1 \cap A_2 \cap A_3).$$
5. Prove that a set A is measurable if and only if for any $\in > 0$ there exist a closed set $B \subseteq A$ such that $m^*(A - B) < \varepsilon$.
6. Show that every non-empty open set has positive measure.
7. The set Q of rationals are enumerated as $x_1, x_2, x_3 \ldots$ and the set G is defined by $G = \bigcup_{n-1}^{\infty}\left]x_n - \frac{1}{n^2}, x_n + \frac{1}{n^2}\right[$.
   Prove that for any closed set F, $m(G\,\Delta F) > 0$.
8. Show that every subset of a set of measure zero is of measure zero.
9. Show that every set with positive outer measure contains a non-measurable subset.
10. If each point of a bounded, measurable set A is covered by a collection I of intervals having arbitrarily small length (called a Vitali covering), then there exists a denumerable set of mutually disjoint intervals $I_1, I_2, \ldots$ such that $\bigcup_{k=1}^{\infty} I_k$ covers A except for a set of measure zero, *i.e.*, $m^*\left(A - \bigcup_{k=1}^{\infty} I_k\right) = 0$ for any set of mutually disjoint intervals $I_1, I_2, \ldots$ which cover A a.e.

11. Define f on $\left[0, \frac{1}{\pi}\right]$ by $f(0) = 0$, $f(x) = x \sin \frac{1}{x}$, $x > 0$. Find the measure of the set $\{x : f(x) \geq 0\}$.
12. Give an example of a measurable set which is not a Borel set.
13. Any function defined on a set of measure zero is measurable.
14. If f_1 and f_2 are measurable functions on any measurable set $A \subseteq [a, b]$ then $\max(f_1, f_2)$ and $\min(f_1, f_2)$ are also measurable.
15. Give an example of a function f such that $|f|$ is measurable, but f is not.
16. If $\{f_n\}$ is a sequence of measurable functions on a measurable set A then $\overline{\lim\limits_{n \to \infty}} f_n$ and $\underline{\lim\limits_{n \to \infty}} f_n$ are measurable on A.
17. Show that the monotone convergence theorem need not hold for decreasing sequence of functions.
18. If $\{f_n\}$ is a sequence of measurable function on a measurable set A and g be an integrable function on A such that $|f_n(x)| \leq g(x)$ a.e. on A, then

$$\int_A \underline{\lim}\, f_n \leq \underline{\lim} \int_A f_n \leq \overline{\lim} \int_A f_n \leq \int_A \overline{\lim}\, f_n .$$

19. Let $f_n(x) = \begin{cases} \frac{x}{n^2} & 0 < x < n \\ 0, & \text{otherwise} \end{cases}$. Evaluate $\operatorname*{Lt}_{n \to \infty} \int_0^n f_n(x)\, dx$ and $\int_0^n \operatorname*{Lt}_{n \to \infty} f_n(x)\, dx$

Are these equal?

20. Evaluate $\int_{-1}^{1} \frac{dx}{x^2 + 1}$.
21. Show that

$$\lim_{n \to \infty} \int_{-n}^{n} \left(1 + \frac{x}{n}\right)^n \exp\left(-x^2/2\right) dx = \sqrt{\frac{\pi}{2e}}.$$

APPENDIX I

Beta and Gamma Functions

We have already discussed the convergence of the improper integrals

$$\int_0^1 x^{m-1} (1-x)^{n-1}\, dx, \text{ and } \int_0^\infty x^{m-1} e^{-x}\, dx,$$

for $m-1<0$, and $n-1<0$ in chapter 11 Sections 3.4 and 4.4 respectively.

We have seen that the first integral converges if $m>0$, $n>0$ and the second converges for $m>0$.

These integrals are named as Beta and Gamma functions, respectively and denoted by $\beta(m, n)$ and $\Gamma(m)$, respectively.

i.e., $$\beta(m, n) = \int_0^1 x^{m-1} (1-x)^{n-1}\, dx = 2\int_0^{\pi/2} \sin^{2m-1}\theta \cos^{2n-1}\theta\, d\theta, \ (x = \sin^2\theta)$$

and

$$\Gamma(m) = \int_0^\infty x^{m-1} e^{-x}\, dx = 2\int_0^\infty r^{2m-1} e^{-r^2}\, dr \quad (x = r^2)$$

Also in chapter 17, example 25, we have established a relation between them, viz.,

$$\beta(m, n) = \frac{\Gamma(m)\,\Gamma(n)}{\Gamma(m+n)}$$

We shall now give *Legendre's Duplication Formula*

$$\sqrt{\pi}\,\Gamma(2m) = 2^{2m-1}\,\Gamma(m)\,\Gamma(m+\frac{1}{2})$$

we have

$$\frac{\Gamma(m)\,\Gamma(n)}{\Gamma(m+n)} = \beta(m, n) = 2\int_0^{\pi/2} \sin^{2m-1}\theta \cos^{2n-1}\theta\, d\theta \qquad \text{...(1)}$$

Taking $n = m$, we have

$$\frac{(\Gamma(m))^2}{\Gamma(2m)} = \beta(m, m) = 2\int_0^{\pi/2} \sin^{2m-1}\theta \cos^{2m-1}\theta\, d\theta$$

$$= \frac{1}{2^{2m-2}}\int_0^{\pi/2} \sin^{2m-1} 2\theta\, d\theta$$

$$= \frac{1}{2^{2m-1}}\int_0^{\pi} \sin^{2m-1}\phi\, d\phi \quad (2\theta = \phi) \qquad \text{...(2)}$$

In (1) taking $n = \dfrac{1}{2}$, we get

$$\frac{\Gamma(m)\,\Gamma(1/2)}{\Gamma(m+\frac{1}{2})} = 2\int_0^{\pi/2} \sin^{2m-1}\theta\, d\theta \qquad \text{...(3)}$$

From (2) and (3), we obtain

$$\frac{(\Gamma(m))^2}{\Gamma(2m)} = \frac{1}{2^{2m-1}}\,\frac{\Gamma(m)\,\Gamma(\frac{1}{2})}{\Gamma(m+\frac{1}{2})}$$

or

$$2^{2m-1}\Gamma(m)\Gamma\left(m+\frac{1}{2}\right) = \sqrt{\pi}\,\Gamma(2m), \text{ since } \Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}.$$

This proves the required duplication formula.

Ex. Prove that $\Gamma(1/4)\,\Gamma(3/4) = \sqrt{2}\pi$.

Next, employing integrating by parts to the Gamma function

$$\Gamma(m) = \int_0^\infty x^{m-1} e^{-x}\, dx, \quad m > 0$$

we obtain

$$\begin{aligned}\Gamma(m+1) &= \lim_{\substack{b\to\infty \\ a\to 0+}} \int_a^b x^m e^{-x}\, dx \\ &= \lim_{\substack{b\to\infty \\ a\to 0+}} \left\{-b^m e^{-b} + a^m e^{-a} + \int_a^b m x^{m-1} e^{-x}\, dx\right\} \\ &= m\,\Gamma(m), \text{ since } b^m e^{-b} \to 0, \text{ as } b \to \infty, \text{ and} \\ &\quad\; a^m e^{-a} \to 0, \text{ as } a \to 0+ \qquad (\because\ m > 0)\end{aligned}$$

$$\therefore \qquad \Gamma(m+1) = m\Gamma(m), \quad \forall\, m > 0.$$

Further, since $\Gamma(1) = 1$, so it can be easily shown that

$$\Gamma(n+1) = n!, \quad \forall\, n \in \mathbf{N}.$$

Ex. 1. Show that

$$\left\{\int_0^{\pi/2} \sin^p x\, dx\right\}\left\{\int_0^{\pi/2} \sin^{p+1} x\, dx\right\} = \frac{\pi}{2(p+1)}$$

Ex. 2. Show that

$$\Gamma(m)\,\Gamma(1-m) = \pi/\sin m\pi, \quad 0 < m < 1$$

[*Hint:* $\beta(m, 1-m) = \dfrac{\Gamma(m)\,\Gamma(1-m)}{\Gamma(1)} = \Gamma(m)\,\Gamma(1-m)$,

and

$$\beta(m, 1-m) = \int_0^1 x^{m-1} (1-x)^{-m}\, dx$$

$$= \int_0^\infty \frac{y^{m-1}}{1+y}\, dy, \text{ taking } x = y/(1+y).$$

Evaluate this improper integral, and use exercise 8, chapter 14.]

Example 1. Show that

$$\int_0^\infty \frac{x^{m-1}}{(1+x)^{m+n}}\, dx = \beta(m, n), \text{ for } m, n > 0$$

■ Put $$x = \frac{t}{1-t}$$

$$dx = \frac{dt}{(1-t)^2}, \text{ when } x \text{ varies from 0 to } \infty, t \text{ varies from 0 to 1.}$$

$$\therefore \quad \int_0^\infty \frac{x^{m-1}}{(1+x)^{m+n}}\, dx = \int_0^1 \left(\frac{t}{1-t}\right)^{m-1} \frac{(1-t)^{m+n}}{(1-t)^2}\, dt$$

$$= \int_0^1 t^{m-1} (1-t)^{n-1}\, dt = \beta(m, n)$$

Example 2. Show that for $l > 0$, $m > 0$

$$\int_a^b (x-a)^{l-1} (b-x)^{m-1}\, dx = (b-a)^{l+m-1}\, \beta(l, m)$$

■ Put $x = py + q$, where p and q are such that when $x = a$, $y = 0$ and when $x = b$, $y = 1$. This gives $p = b - a$, and $a = a$ and the integral becomes

$$= \int_0^1 [(b-a)\, y + a - a]^{l-1}\, (b - (b-a)\, y - a)^{m-1}\, (b-a)\, dy$$

$$= \int_0^1 (b-a)^{l-1+1+m-1}\, y^{l-1}\, (1-y)^{m-1}\, dy = (b-a)^{l+m-1}\ \beta(l, m)$$

Example 3. Show that

$$\int_0^1 x^{-1/3} (1-x)^{-2/3} (1+2x)^{-1}\, dx = \frac{1}{9^{1/3}}\ \beta(2/3, 1/3)$$

■ Put $\dfrac{x}{1-x} = \dfrac{at}{1-t}$, where a is a constant to be chosen so that the given integral becomes Beta function

$$x = \frac{at}{1-(1-a)t}\ .$$

$$dx = \frac{2dt}{[1-(1-a)t]^2} \text{ when } x = 0,\ t = 0$$

$$\int_0^1 \left[\frac{at}{1-(1-a)t}\right]^{-1/3} \left[\frac{1-t}{1-(1-a)t}\right]^{-2/3} \left[\frac{1-t+3at}{1-(1-a)t}\right]^{-1} \frac{adt}{[1-(1-a)t]^2}$$

$$= \int_0^1 \frac{a^{2/3}\, t^{-1/3}\, [1-t(1-3a)]^{-1}\, dt}{(1-t)^{2/3}}\ .$$

If we choose $a = \frac{1}{3}$ then the integral becomes a Beta function and therefore taking $a = \frac{1}{3}$, we have

$$-\int_0^1 \left(\frac{1}{3}\right)^{2/3} t^{(2/3)-1}\, (1-t)^{(1/3)-1}\, dt = \frac{1}{9^{1/3}}\, \beta(2/3, 1/3)$$

Example 4. If n is a positive integer, prove that the ratio of the areas enclosed by the curves

$$x^{2n} + y^2 = 1,\ x^{2n} + y^{2n} = 1 \text{ is } n2^{1/n}/(n+1)$$

■ For area under the 1st curve

Put $x^{2n} = \cos^2\theta,\ y^2 = \sin^2\theta$

then the area is

$$A_1 = 4\int_0^{\pi/2} \sin\theta\, \frac{1}{n}\cos^{(1/n)-1}\theta\,(-\sin\theta)\, d\theta$$

$$= -\frac{4}{n}\int_0^{\pi/2} \sin^2\theta\, \cos^{(1/n)-1}\theta\, d\theta$$

$$= -\frac{2}{n}\beta\left(\frac{3}{2}, \frac{1}{2n}\right) = -\frac{2}{n}\frac{\Gamma(3/2)\,\Gamma(1/2n)}{\Gamma(1/2n + 3/2)}$$

Similarly putting $x^{2n} = \cos^2\theta,\ y^{2n} = \sin^2\theta,$ the area under the 2nd curve is

$$A_2 = -4\int_0^{\pi/2} \sin^{1/n}\theta\ \ \frac{1}{n}\cos^{(1/n)-1}\theta \sin\theta\, d\theta$$

$$= -\frac{4}{n}\int_0^{\pi/2} \sin^{(1/n)+1}\theta\, \cos^{(1/n)-1}\theta d\theta$$

$$= -\frac{2}{n}\beta\left(\frac{1}{2n} + 1, \frac{1}{2n}\right)$$

$$\therefore \qquad \frac{A_1}{A_2} = \frac{\frac{1}{2}\Gamma\left(\frac{1}{2}\right)\Gamma\left(\frac{1}{2n}\right)\Gamma\left(\frac{1}{2n}+1+\frac{1}{2n}\right)}{\Gamma\left(\frac{1}{2n}+\frac{1}{2}+1\right)\Gamma\left(\frac{1}{2n}+1\right)\Gamma\left(\frac{1}{2n}\right)}$$

$$= \frac{1}{2} \frac{\sqrt{\pi}\, \Gamma\left(\frac{1}{2n}\right) \frac{1}{n} \Gamma\left(\frac{1}{n}\right)}{\left(\frac{1}{2n} + \frac{1}{2}\right) \Gamma\left(\frac{1}{2n} + \frac{1}{2}\right) \frac{1}{2n} \Gamma\left(\frac{1}{2n}\right) \Gamma\left(\frac{1}{2n}\right)}.$$

$$= \frac{2n}{n+1} \frac{\Gamma\left(\frac{1}{n}\right) \sqrt{\pi}}{\Gamma\left(\frac{1}{2n}\right) \Gamma\left(\frac{1}{2n} + \frac{1}{2}\right)}$$

Using duplication formula,

$$2^{1/n-1}\, \Gamma\left(\frac{1}{2n}\right) \Gamma\left(\frac{1}{2n} + \frac{1}{2}\right) = \sqrt{\pi}\, \Gamma\left(\frac{1}{n}\right)$$

we get

$$\frac{A_1}{A_2} = \frac{2n \sqrt{\pi}\, \Gamma\left(\frac{1}{n}\right)}{(n+1)\, 2^{-1/n+1} \sqrt{\pi}\, \Gamma\left(\frac{1}{n}\right)} = 2^{1/n} \frac{n}{n+1}.$$

Example 5. Evaluate the integrals

$$\int_0^\infty e^{-ax} x^{m-1} \cos bx\, dx, \text{ and } \int_0^\infty e^{-ax} x^{m-1} \sin bx\, dx, \quad m > 0.$$

Hence or otherwise show that

$$\int_0^\infty x^{m-1} \cos bx\, dx = \frac{\Gamma(m)}{b^m} \cos\left(\frac{m\pi}{2}\right) \text{ and } \int_0^\infty x^{m-1} \sin bx\, dx = \frac{\Gamma(m)}{b^m} \sin(m\pi/2).$$

■ Now $$\int_0^\infty e^{-kx} x^{m-1}\, dx = \frac{\Gamma(m)}{k^m}$$

Taking $k = a - ib$, $|k| > 0$

$$\int_0^\infty e^{-(a-ib)x} x^{m-1}\, dx = \frac{\Gamma(m)}{(a-ib)^m}$$

$$\int_0^\infty e^{-ax} e^{ibx} x^{m-1}\, dx = \frac{\Gamma(m)\, (a+ib)^m}{(a-ib)^m (a+ib)^m}$$

$$\int_0^\infty e^{-ax} (\cos bx + i \sin bx) x^{m-1}\, dx = \frac{\Gamma(m)\, (a+ib)^m}{(a^2+b^2)^m}$$

Writing $a + ib = r(\cos\theta + i\sin\theta)$, and separating the real and imaginary parts, we get

$$\int_0^\infty e^{-ax}\cos bx\, x^{m-1}\,dx = \frac{\Gamma(m)\cos m\theta}{(a^2+b^2)^{m/2}},\ \text{where } \theta = \tan^{-1}\frac{b}{a}$$

and

$$\int_0^\infty e^{-ax}\sin bx\, x^{m-1}\,dx = \frac{\Gamma(m)\sin m\theta}{(a^2+b^2)^{m/2}},\ \text{where } \theta = \tan^{-1}\frac{b}{a}$$

Taking $a = 0, \theta = \pi/2$

$$\int_0^\infty \cos bx\, x^{m-1}\,dx = \frac{\Gamma(m)\cos(m\pi/2)}{b^m}$$

and

$$\int_0^\infty \sin bx\, x^{m-1}\,dx = \frac{\Gamma(m)\sin(m\pi/2)}{b^m}.$$

EXERCISE

1. Show that

$$\int_0^p x^m (p^q - x^q)^n\,dx = \frac{p^{qn+m+1}}{q}\,\beta\left(n+1, \frac{m+1}{q}\right)$$

if $p > 0, q > 0, m + 1 > 0, n + 1 > 0$.

2. Prove that

(i) $\displaystyle\int_0^1 \frac{x^{m-1}+x^{n-1}}{(1+x)^{m+n}}\,dx = \beta(m,n).$

$$\left[\textit{Hint}:\ ?\,(m,n) = \int_0^\infty \frac{x^{m-1}\,dx}{(1+x)^{m+n}} = \int_0^1 \frac{x^{m-1}}{(1+x)^{m+n}}\,dx + \int_1^\infty \frac{x^{m-1}}{(1+x)^{m+n}}\,dx\right.$$

$$\left.\text{Put } x = \frac{1}{t} \text{ in the second integral.}\right]$$

(ii) $\displaystyle\int_0^\infty \frac{(x^{m-1}+x^{n-1})}{(1+x)^{m+n}}\,dx = 2\beta(m,n).$

3. Show that for $m, n > 0$,

$$\int_0^1 \frac{x^{m-1}(1-x)^{n-1})}{(b+cx)^{m+n}}\,dx = \frac{\beta(m,n)}{(b+c)^m\, b^n}.$$

$$\left[\textit{Hint}:\ \text{Put } y = \frac{(b+c)x}{b+cx}\right]$$

4. Show that

$$\int_0^{\pi}\left(1-\frac{t}{n}\right)^n t^{x-1}\,dt = n^x\,\beta(x, n+1),\text{ where } x>0.$$

5. Show that for $m > 0$,

(*i*) $\beta(m, m) = 2^{1-2m}\,\beta(m, \tfrac{1}{2})$,

(*ii*) $\beta(m, m)\,\beta(m+\tfrac{1}{2}, m+\tfrac{1}{2}) = \pi m^{-1} 2^{1-4m}$.

6. Prove that

(*i*) $\displaystyle\int_0^{\pi/2}\frac{d\theta}{\sqrt{\sin\theta}}\times\int_0^{\pi/2}\sqrt{\sin\theta}\,d\theta = \pi.$

(*ii*) $\displaystyle\left\{\int_0^1\frac{x^2dx}{\sqrt{1-x^4}}\right\}\left\{\int_0^1\frac{dx}{\sqrt{1+x^4}}\right\} = \frac{\pi}{4\sqrt{2}}.$

7. Show that

(*i*) $\displaystyle\int_0^1\sqrt{1-x^4}\,dx = \frac{1}{12}\sqrt{\frac{2}{\pi}}\,[\Gamma(\tfrac{1}{4})]^2,$

(*ii*) $\displaystyle\int_0^1(1-x^n)^{-1/2}\,dx = 2^{(2/n)-1}\,[\Gamma(1/n)]^2/n\Gamma(2/n).$

8. Show that the perimeter of the lemniscate $r^2 = 2a^2\cos 2\theta$ is

$$\frac{a}{\sqrt{\pi}}\,[\Gamma(\tfrac{1}{4})]^2.$$

9. Show that the perimeter of a loop of the curve

$$r^n = a^n\cos n\theta$$

is

$$\frac{a}{n}\cdot 2^{\left(\frac{1}{n}\right)-1}\,\frac{\left[\Gamma\left(\frac{1}{2n}\right)\right]^2}{\Gamma\left(\frac{1}{n}\right)}.$$

10. Show that the area bounded by the curve $x^n + y^n = a^n$, and the co-ordinate axes in the first quadrant is $[\Gamma(1/n)]^2/2n\Gamma(2/n)$.

Appendix II

Cantor's Theory of Real Numbers

In chapter 1, we followed Dedekind's theory to extend the rational number system to the real number system. We now introduce another method, which is less algebraic but more sophisticated, due to G. Cantor. The construction of real numbers by Cantor's theory depends on the existence of an equivalence relation in the set of all Cauchy sequences of rational numbers. Defining each equivalence class as a real number and by suitably defining addition and multiplication, and order, this set **R** of Cantor's real numbers will be made into an ordered field which will be an extension of the ordered field **Q** or rational numbers. It will then be shown that this ordered field **R** is also *complete* (in the sense defined earlier) i.e., the field **R** is *order complete*.

Since the concept of *Cauchy sequences of rational numbers* and their limits are basic to Cantor's Theory, we start the discussion by a brief introduction to these sequences.

1. SEQUENCES OF RATIONAL NUMBERS

Definitions

1. A function S on the set **N** of natural numbers into the set **Q** of rational numbers is called a *rational sequence* or a *sequence of rational numbers* and is symbolically denoted as $\mathbf{S} : \mathbf{N} \to \mathbf{Q}$.
2. A rational sequence $\{S_n\}$ is said to be *bounded* if there exists a rational numbers $K > 0$ such that
$$|S_n| \le K, \quad \forall\, n.$$
3. A sequence $\{S_n\}$ of rational numbers is called a *Cauchy sequence* or a *fundamental sequence*, if for each rational number $\varepsilon > 0$ there exists a positive integer m_0, such that
$$|S_n - S_m| < \varepsilon, \quad \forall\, m, n \ge m_0$$
or
$$|S_{n+p} - S_n| < \varepsilon, \quad \forall\, n \ge m_0, \text{ integer } p \ge 1.$$
4. A rational sequence $\{S_n\}$ is said to *converge* to a rational number l (or we have the number l as its *limit*) if for each rational number $\varepsilon > 0$, there exists a positive integer m_0, (depending on ε), such that
$$|S_n - l| < \varepsilon \text{ for all } n \ge m_0, \text{ and we then write}$$
$$\lim_{n\to\infty} S_n = l \text{ or } S_n \to l$$

It may be easily shown that

(*i*) every Cauchy sequence is bounded,

(*ii*) every convergent sequence is bounded, and

(*iii*) every convergent sequence is a Cauchy sequence, i.e., a necessary condition for convergence of a rational sequence is that it is a Cauchy sequence.

Thus *a necessary condition for convergence* of a rational sequence is that for any rational number $\varepsilon > 0$, there exists a positive integer m_0 such that

$$|S_{n+p} - S_n| < \varepsilon, \quad \forall\, n \geq m_0,\ p \geq 1 \qquad \text{...(1)}$$

Example. Show that the sequence $\{S_n\}$, where

$$S_{n+1} = \frac{2 + S_n}{1 + S_n}, \quad n \geq 1$$

$$S_1 = 1$$

is not convergent in the field or rational numbers.

If the sequence converges to a rational number, say q, then

$$\lim S_{n+1} = q = \lim S_n$$

$$\therefore \qquad q = \lim \frac{2 + S_n}{1 + S_n} = \frac{2 + q}{1 + q}$$

or $$q^2 = 2$$

But no rational number exists whose square is equal to 2. Hence, the sequence is not convergent in the field of rational numbers.

Note: $\{S_n\}$ is the rational sequence, $1, \frac{3}{2}, \frac{7}{5}, \frac{17}{12}, \ldots$

Thus by considering the sequence $\{S_n\}$, where $S_{n+1} = \dfrac{2 + S_n}{1 + S_n}$, for $n \geq 1$, $S_1 = 1$, it may be shown that (1) is not a sufficient condition for convergence of a rational sequence. Thus every Cauchy sequence is not convergent in the field of rationals (compare § 6.1, Ch. 3).

Ex. 1. Show that the sequence $\{S_n\}$, where $S_{n+1} = \dfrac{l + 2 + kS_n}{k + lS_n}$, $S_1 = 1$, l, k are finite non-zero numbers, is not convergent in the field of rationals.

Ex. 2. If $\{a_n\}$ and $\{b_n\}$ are Cauchy sequences of rational numbers, then

(*i*) the sequences $\{a_n \pm b_n\}$, $\{a_n b_n\}$ are also Cauchy sequences,

(*ii*) the sequences $\left\{\dfrac{1}{b_n}\right\}$ and $\left\{\dfrac{a_n}{b_n}\right\}$ are also Cauchy sequences provided $\{b_n\}$ does not converge to zero, and $b_n \neq 0$ for any n.

2. CANTOR REAL NUMBER

We shall use F_Q to denote the set of all Cauchy sequences of rational numbers.

Definition. A sequence $\{a_n\} \in F_Q$ is said to be equivalent to $\{b_n\} \in F_Q$ whenever $\{a_n - b_n\}$ converges to zero. Expressed symbolically,

$$\{a_n\} \sim \{b_n\} \text{ iff } \lim (a_n - b_n) = 0 .$$

Theorem 1. *The relation $\sim$ in the set of all Cauchy sequences of rational numbers, defined by*

$$\{a_n\} \sim \{b_n\} \text{ iff } \lim (a_n - b_n) = 0$$

is an equivalence relation.

(1) For any Cauchy sequence $\{a_n\} \in F_Q$

$$\lim (a_n - a_n) = 0$$

$$\therefore \quad \{a_n\} \sim \{a_n\}$$

and so the relation $\sim$ is reflexive.

(2) Let $\{a_n\}, \{b_n\}$ be sequences in F_Q such that $\lim (a_n - b_n) = 0$.
Then

$$\lim (b_n - a_n) = \lim [-(a_n - b_n)] = 0$$

$$\therefore \quad \{a_n\} \sim \{b_n\} \Rightarrow \{b_n\} \sim \{a_n\}$$

and so the relation $\sim$ is symmetric.

(3) Let $\{a_n\}, \{b_n\}, \{c_n\}$ be sequences in F_Q, such that

$$\{a_n\} \sim \{b_n\} \text{ and } \{b_n\} \sim \{c_n\}$$

i.e.,

$$\lim (a_n - b_n) = 0 \text{ and } \lim (b_n - c_n) = 0$$

Then,

$$\lim (a_n - c_n) = \lim (a_n - b_n + b_n - c_n)$$
$$= \lim (a_n - b_n) + \lim (b_n - c_n) = 0$$

$$\therefore \quad \{a_n\} \sim \{b_n\} \wedge \{b_n\} \sim \{c_n\} \Rightarrow \{a_n\} \sim \{c_n\}$$

Hence, the relation $\sim$ is transitive.

Hence from (1), (2) and (3) it follows that the relation $\sim$ is an equivalence relation in F_Q.

Notation. Let $[a_n]$ denote the equivalence class containing the sequence $\{a_n\}$, i.e., the set of rational Cauchy sequences equivalent to $\{a_n\}$. Thus,

$$[a_n] = \{\{x_n\} \in F_Q \mid \{x_n\} \sim \{a_n\}\}.$$

Ex. 1. If $\{a_n\} \in F_Q$, then $\lim (a_n) = a$ iff $\{a_n\} \sim \{a\}$.

Ex. 2. Two equivalence classes $[a_n]$, and $[a_n']$ are equal iff $\{a_n\} \sim \{a_n'\}$.

Definition. A Cantor *real number* is an equivalence class $[a_n]$ with respect to the equivalence relation $\sim$ in F_Q defined by the condition

$$\{a_n\} \sim \{b_n\} \text{ iff } \lim (a_n - b_n) = 0$$

Thus if ξ is a real number $[a_n]$, then

$$\xi = \{\{x_n\} \in F_Q \mid \{x_n\} \sim \{a_n\}\}$$

We shall denote by **R** the set of all real numbers and use $\xi, \eta, \ldots$ to denote the real numbers.

3. ADDITION AND MULTIPLICATION IN R

We shall define two binary operations (+ and .) in **R**, to be called addition and multiplication and discuss some of their properties. But we shall need the following theorems before we can do so.

Theorem 2. *If $\{x_n\}$, $\{y_n\}$, $\{a_n\}$, and $\{b_n\}$ belong to F_Q such that $\{x_n\} \sim \{a_n\}$ and $\{y_n\} \sim \{b_n\}$, then $\{x_n + y_n\}$ and $\{a_n + b_n\}$ also belong to F_Q and $\{x_n + y_n\} \sim \{a_n + b_n\}$.*

Since $\{x_n\}$, $\{y_n\}$, $\{a_n\}$ and $\{b_n\}$ all belong to F_Q therefore as in Ex. 2 § 1, $\{x_n + y_n\}$ and $\{a_n + b_n\}$ also belong to F_Q. Also

$$\{x_n\} \sim \{a_n\} \Rightarrow \lim (x_n - a_n) = 0$$

and

$$\{y_n\} \sim \{b_n\} \Rightarrow \lim (y_n - b_n) = 0$$

$$\therefore \quad \lim (x_n + y_n - a_n - b_n) = \lim (x_n - a_n) + \lim (y_n - b_n) = 0$$

$$\Rightarrow \quad \{x_n + y_n\} \sim \{a_n + b_n\}.$$

Theorem 3. *If $\{x_n\}$, $\{y_n\}$, $\{a_n\}$ and $\{b_n\}$ belong to F_Q such that $\{x_n\} \sim \{a_n\}$ and $\{y_n\} \sim \{b_n\}$, then $\{x_n y_n\}$ and $\{a_n b_n\}$ also belong to F_Q and $\{x_n y_n\} \sim \{a_n b_n\}$.*

Since $\{x_n\}$, $\{y_n\}$, $\{a_n\}$ and $\{b_n\}$ all belong to F_Q, therefore $\{x_n y_n\}$ and $\{a_n b_n\}$ also belong to F_Q.

Again since $\{x_n\}$, $\{b_n\}$ are rational Cauchy sequences, they are bounded and therefore $\exists$ positive rational numbers k_1, k_2 such that

$$|x_n| < k_1, |b_n| < k_2, \quad \forall\, n \in \mathbf{N}$$

Also since $\{x_n\} \sim \{a_n\}$ and $\{y_n\} \sim \{b_n\}$,

$$\therefore \quad \lim (x_n - a_n) = 0 \text{ and } \lim (y_n - b_n) = 0$$

so that for rational $\varepsilon > 0$, $\exists$ positive integers m_1, m_2 that

$$|x_n - a_n| < \varepsilon/2k_2, \quad \text{for } n \geq m_1$$

and

$$|y_n - b_n| < \varepsilon/2k_1, \quad \text{for } n \geq m_2$$

Let $m = \max(m_1, m_2)$.

Hence for $n \geq m$, we have

$$\begin{aligned} |x_n y_n - a_n b_n| &= |x_n(y_n - b_n) + b_n(x_n - a_n)| \\ &\leq |x_n||y_n - b_n| + |b_n||x_n - a_n| \\ &< k_1 \frac{\varepsilon}{2k_1} + k_2 \frac{\varepsilon}{2k_2} = \varepsilon \end{aligned}$$

$\Rightarrow \qquad \lim (x_n y_n - a_n b_n) = 0$

$\therefore \qquad \{x_n y_n\} \sim \{a_n b_n\}$

Theorem 4. (*Addition in* **R**). *There is a binary operation* $f: \mathbf{R} \times \mathbf{R} \to \mathbf{R}$ *such that for every pair of real numbers* $\xi = [a_n]$, $\eta = [b_n]$,

$$f(\xi, \eta) = \zeta$$

where $\zeta = [a_n + b_n]$

Let (ξ, η) be an arbitrary element of $\mathbf{R} \times \mathbf{R}$ and let $\xi = [a_n]$, $\eta = [b_n]$, for some sequences $\{a_n\}$ and $\{b_n\}$ in F_Q.

Since $\{a_n + b_n\} \in F_Q$, therefore $[a_n + b_n] \in \mathbf{R}$.

Let $f(\xi, \eta) = \zeta$, where $\zeta = [a_n + b_n]$.

We now show that mapping f is well defined.

Let, if possible, $f(\xi, \eta) = \zeta'$, where $\zeta' = [a_n' + b_n']$, $\xi = [a_n']$, and $\eta = [b_n']$.

Now

$$[a_n] = \xi = [a'_n] \text{ and } \eta = [b_n] = [b_n']$$

$\therefore \qquad \{a_n\} \sim \{a_n'\}$ and $\{b_n\} \sim \{b_n'\}$

$\Rightarrow \qquad \{a_n + b_n\} \sim \{a_n' + b_n'\}$

$\Rightarrow \qquad [a_n + b_n] = [a_n' + b_n']$

$\Rightarrow \qquad \zeta = \zeta'$

Thus f is a binary operation on **R**.

Definition. The binary operation f on **R** is called *addition* in **R** and is denoted by +. Thus

For every pair ξ, η *of real numbers where* $\xi = [a_n]$ *and* $\eta = [b_n]$, *for some sequences* $\{a_n\}$ *and* $\{b_n\}$ *in* F_Q, *the real number* $\zeta = [a_n + b_n]$ *is called the sum of* ξ *and* η *and is denoted by* $\xi + \eta$.

Ex. For real numbers $\xi, \eta, \zeta, \ldots$, prove that

(*i*) $\xi + \eta = \eta + \xi$

(*ii*) $\xi + (\eta + \zeta) = (\xi + \eta) + \zeta$

(*iii*) There is a unique real number 0 such that

$$\xi + 0 = 0 + \xi = \xi$$

[*Hint*: Take $0 = [0_n]$, where $0_n = 0$ for all n]

(*iv*) For every $\xi \in \mathbf{R}$, there exists a unique $\eta \in \mathbf{R}$ such that

$$\xi + \eta = 0 = \eta + \xi.$$

Theorem 5. (*Multiplication in* **R**). *There is a binary operation* $g: \mathbf{R} \times \mathbf{R} \to \mathbf{R}$ *such that for every pair of real numbers* $\xi = [a_n], \eta = [b_n]$,

$$g(\xi, \eta) = \zeta,$$

where $\zeta = [a_n b_n]$.

Let (ξ, η) be an arbitrary element of $\mathbf{R} \times \mathbf{R}$ and let $\xi = [a_n], \eta = [b_n]$, for some sequences $\{a_n\}$ and $\{b_n\}$ in F_Q.

Since $\{a_n b_n\} \in F_Q$, therefore $[a_n b_n] \in \mathbf{R}$.

Let $g(\xi, \eta) = \zeta$, where $\zeta = [a_n b_n]$.

We now show that the mapping g is well defined.

Let, if possible, $g(\xi, \eta) = \zeta'$, where $\zeta' = [a_n' b_n']$, and $\xi = [a_n'], \eta = [b_n']$.

Now

$$[a_n] = [a_n'], \text{ and } [b_n] = [b_n']$$

$$\therefore \quad \{a_n\} \sim \{a_n'\} \text{ and } \{b_n\} \sim \{b_n'\}$$

$$\Rightarrow \quad \{a_n b_n\} \sim \{a_n' b_n'\}$$

$$\Rightarrow \quad [a_n b_n] = [a_n' b_n']$$

$$\Rightarrow \quad \zeta = \zeta'$$

Thus g is a binary operation on **R**.

Definition. The binary operation g on **R** is called *multiplication* in **R** and is denoted by (.). Thus for every pair ξ, η or real numbers where $\xi = [a_n]$ and $\eta = [b_n]$ for some sequences $\{a_n\}$ and $\{b_n\}$ in F_Q, the real number $\zeta = [a_n b_n]$ is called the product of ξ and η and is denoted by $\xi \cdot \eta$.

Ex. For real numbers $\xi, \eta, \zeta, \ldots$, show that

(*i*) $\xi \cdot \eta = \eta \cdot \xi$

(*ii*) $\xi \cdot (\eta \cdot \zeta) = (\xi \cdot \eta) \cdot \zeta$,

(*iii*) There exists a real number 1, called the multiplicative identity such that

$$\xi \cdot 1 = 1 \cdot \xi = \xi$$

[*Hint*: Take $1 = [I_n]$, where $I_n = 1$ for all n.]

(*iv*) The multiplicative identity is unique.

(v) To each real number $\xi \neq 0$, there corresponds a unique real number η such that

$$\xi \cdot \eta = 1 = \eta \cdot \xi$$

Let $\xi = [a_n]$, Since $\xi \neq 0$, the sequence $\{a_n\}$ can have at the most a finite number of terms equal to zero. Let $m \in \mathbf{N}$ be such that $a_n \neq 0$ for $n > m$.

Let us define a sequence $\{b_n\}$, where $b_n = \dfrac{1}{a_n}$, for $n > m$, and $b_n = 1$ for $n \leq m$.

The real number $\eta = [b_n]$ has the desired property.

The real number η is called the *multiplicative inverse* of ξ and is denoted by ξ^{-1}.

(vi) The multiplicative inverse ξ^{-1} is unique.

Theorem 6. *Prove that* $(R, +, .)$ *is a field.*

The proof is left to the reader.

4. ORDER IN R

We shall now give an *order structure* to the field of real numbers. To do so we shall first define the set of positive elements of **R**.

4.1 *Definition (A positive sequence of rational numbers). A rational sequence* $\{a_n\}$ *is called a positive sequence if there exists a positive rational number e and a positive integer m such that*

$$a_n > e, \text{ for all } n \geq m$$

From the above definition it may be easily shown that *if* $\{a_n\}$ *and* $\{b_n\}$ *are positive sequences of rational numbers, then* $\{a_n + b_n\}$ *and* $\{a_n b_n\}$ *are also positive sequences of rational numbers*.

Theorem 7. *If* $\{a_n\} \sim \{a_n'\}$, *and* $\{a_n\}$ *is a positive rational sequence, then* $\{a_n'\}$ *is also a positive rational sequence*.

Since $\{a_n\}$ is a positive rational sequence, therefore $\exists\ e > 0$ in $\mathbf{Q}^+$ and $m_1 \in \mathbf{N}$ such that

$$a_n > e, \text{ for } n \geq m_1 \qquad \ldots(1)$$

Again since $\{a_n\} \sim \{a_n'\}$, therefore $\exists\ m_2 \in \mathbf{N}$ such that

$$|a_n' - a_n| < \frac{e}{2}, \text{ for all } n \geq m_2$$

or

$$-\frac{e}{2} < a_n' - a_n < \frac{e}{2}, \text{ for } n \geq m_2 \qquad \ldots(2)$$

Let $m_0 = \max(m_1, m_2)$.

$$\therefore \quad a_n' = (a_n' - a_n) + a_n$$

$$> -\frac{e}{2} + e, \text{ for } n \geq m_0 \qquad \text{[from (1) and (2)]}$$

$$= \frac{e}{2} > 0, \text{ for } all\ n \geq m_0$$

Hence $\{a_n'\}$ is a positive rational sequence.

Corollary. If ξ is a real number and $\{a_n\} \in \xi$ be a positive sequence in Q then every sequence $\{a_n'\} \in \xi$ is also a positive sequence in **Q**.

4.2 *Definition. A real number ξ is positive if every sequence in ξ is a positive sequence.*

In view of Theorem 7.1, it follows that a *real number ξ is positive if and only if there exists a positive rational sequence in ξ.*

We shall denote the set of positive real numbers by $\mathbf{R}^+$. Thus

$$\mathbf{R}^+ = \begin{cases} \{\xi \in \mathbf{R} \mid \xi \text{ is positive}\} \\ \{\xi \in \mathbf{R} \mid \text{for some } \{a_n\} \in \xi, \{a_n\} \text{ is positive}\} \end{cases}$$

From definition, the following result may be easily proved.

If ξ, η *are positive real numbers then so also are* $\xi + \eta$ and $\xi \cdot \eta$.

Theorem 8. *If $\xi \in \mathbf{R}$, then one and only one of the following statements is true:*

$(i)\ \xi = 0, \quad (ii)\ \xi \in \mathbf{R}^+, \quad (iii)\ -\xi \in \mathbf{R}^+.$

We first show that at *least one* of the three statements is true.

Let $\xi = [a_n]$, so that $\{a_n\} \in \xi$, and $\{a_n\} \in F_Q$.

When $\xi = 0$, there is nothing to prove, for then (i) holds.

Let $\xi \neq 0$, so that $\{a_n\}$ is not equivalent to $\{0_n\}$, where $0_n = 0$ for all n, consequently $\{a_n\}$ does not converge to 0. Hence $\exists\, e \in \mathrm{Q}^+$ and $m_1 \in \mathbf{N}$ such that

$$|a_n| > e \quad \forall\ n \geq m_1 \qquad ...(1)$$

Again since $\{a_n\}$ is a Cauchy sequence and $e > 0$, therefore $\exists\, m_2 \in \mathbf{N}$ such that

$$|a_{n+p} - a_n| < \frac{e}{2}, \text{ for } n \geq m_2,\ p \geq 1 \qquad ...(2)$$

Let $m = \max\,(m_1, m_2)$, then from (2),

$$a_m - \frac{e}{2} < a_{m+p} < a_m + \frac{e}{2},\ p \geq 1 \qquad ...(3)$$

From (1) either $a_m > e$ or $a_m < -e$.

If $a_m > e$, then from (3),

$$a_{m+p} > e - \frac{e}{2} = \frac{e}{2} > 0,\ p \geq 1$$

Therefore $\{a_n\}$ is a positive sequence and hence $\xi \in \mathbf{R}^+$.

And if $a_m < -e$, then again from (3),

$$a_{m+p} < -e + \frac{e}{2} = -\frac{e}{2},\ p \geq 1$$

i.e.,
$$-a_{m+p} > \frac{e}{2} > 0,\ p \geq 1$$

Therefore $\{-a_n\}$ is a positive sequence and hence $-\xi \in \mathbf{R}^+$.

Thus we have shown that at least one of the three statements is true.

We now proceed to show that *not more than one* of the three statements is true.

If $\xi = 0$, then $\{a_n\} \sim \{0_n\}$. Hence for rational $e > 0, \exists\, m' \in \mathbf{N}$ such that

$$|a_n| < e, \text{ for } n \geq m'$$

Hence there is no $e \in \mathbf{Q}^+$ such that for some $m_0 \in \mathbf{N}$, either

$$a_n \geq e \qquad \forall\, n \geq m_0$$

or
$$-a_n \geq e \qquad \forall\, n \geq m_0$$

Thus if $\xi = 0$, then neither $\xi \in \mathbf{R}^+$ nor $-\xi \in \mathbf{R}^+$.

Now, if possible, $\xi \in \mathbf{R}^+$ and $-\xi \in \mathbf{R}^+$.

Then for some e, e' $(e < e'$, say$) \in \mathbf{Q}^+$ and $n_1, n_2 \in \mathbf{N}$,

$$a_n \geq e \qquad \forall\, n \geq n_1$$
$$-a_n \geq e' \qquad \forall\, n \geq n_2$$

Hence for $n = \max(n_1, n_2)$,

$$0 < e' \leq -a_n \leq -e < 0$$

which is impossible,

$\therefore$
$$\xi \in \mathbf{R}^+ \ \wedge\ -\xi \in \mathbf{R}^+ \text{ is false}$$

Hence the theorem.

4.3 *Thus $\mathbf{R}^+$ is a set of positive elements of $\mathbf{R}$ such that*

(*a*) *If $\xi, \eta \in \mathbf{R}^+$, then $\xi + \eta \in \mathbf{R}^+$ and $\xi . \eta \in \mathbf{R}^+$.*

(*b*) *For each $\xi \in \mathbf{R}$, one and only one of the following is true:*

(*i*) $\xi \in \mathbf{R}^+$ (*ii*) $\xi = 0$ (*iii*) $-\xi \in \mathbf{R}^+$.

4.4 *Definition. A real number ξ is said to be greater than a real number η* if $\xi - \eta \in \mathbf{R}^+$.

Using the symbol > to denote '*greater than*', we write

$$\xi > \eta \text{ iff } \xi - \eta \in \mathbf{R}^+$$

The same thing can also be expressed by saying that η is '*less than*' ξ and write $\eta < \xi$.

Ex. For real numbers $\xi, \eta, \zeta, \ldots,$ prove that

1. Exactly one of the following holds:

   $(i)\ \xi > \eta, \quad (ii)\ \xi = \eta, \quad (iii)\ \xi < \eta.$

2. $\xi > \eta$ and $\eta > \zeta \Rightarrow \xi > \zeta.$
3. $\xi > \eta \Rightarrow \xi + \zeta > \eta + \zeta.$
4. For $\zeta \in \mathbf{R}^+, \xi > \eta \Rightarrow \xi \cdot \zeta > \eta \cdot \zeta.$

Theorem 9. *Prove that* $(\mathbf{R}, +, \cdot, >)$ *is an ordered field.*

The proof is left to the reader.

5. REAL RATIONAL AND IRRATIONAL NUMBERS

If a Cauchy sequence of rational numbers converges in the field of rationals then the real number determined by it is called a *real rational number*. Thus if a rational Cauchy sequence $\{a_n\}$ converges to a rational number α then $[a_n]$ is the real rational number α.

Theorem 10. *If α is a rational number then there exists a rational Cauchy sequence converging to α.*

Let $\{a_n\}$, where $a_n = \alpha$ for all n, be a rational sequence, which clearly converges to α. Also it is a Cauchy sequence, since $|a_{n+p} - a_n| = 0$ for all n, p.

Theorem 11. *If $\{a_n\}$ and $\{b_n\}$ are two rational Cauchy sequences converging to the same limit α, then*

$$\{a_n\} \sim \{b_n\}$$

$$\because \quad \lim a_n = \alpha = \lim b_n$$

$$\therefore \quad \lim (a_n - b_n) = 0$$

$$\therefore \quad \{a_n\} \sim \{b_n\}$$

Theorem 12. *To every rational number α there corresponds a unique real rational number.*

Let $\{a_n\}$ be a rational Cauchy sequence that converges to α, so that $\{a_n\}$ is a real rational number that corresponds to α.

If $[a_n']$ be another real rational number that corresponds to α, then the rational Cauchy sequence $\{a_n'\}$ also converges to α. Hence by Theorem 11,

$$\{a_n\} \sim \{a_n'\}$$

$$\therefore \quad [a_n] = [a_n'].$$

Notation. The set of all real rational numbers will be denoted by $\mathbf{R}^*$, where $\mathbf{R}^* = \{\alpha \mid \alpha$ is a real rational number$\}$.

If $\xi \in \mathbf{R}$ but $\xi \notin \mathbf{R}^*$, then ξ is called *real irrational* number and the set of all real irrational numbers is $\mathbf{R} - \mathbf{R}^*$,

6. SOME PROPERTIES OF REAL NUMBERS

Theorem 13. *If ξ is a real number and $\{x_n\}$ a Cauchy sequence in* $\mathbf{Q}$ *such that $\{x_n\} \in \xi$, then $lim\ x_n = \xi$* in $\mathbf{R}$.

Let $\varepsilon > 0$ be a real number and e a rational number such that $0 < e < \varepsilon$.

Since $\{x_n\}$ is a Cauchy sequence in $\mathbf{Q}$, therefore for $e > 0$, $\exists\, m_0 \in \mathbf{N}$ such that

$$|x_n - x_m| < \frac{e}{2} \quad \forall\, m, n \geq m_0$$

Hence for all $m, n \geq m_0$, $\quad e - |x_n - x_m| > \dfrac{e}{2}$

and so for each $n \geq m_0$ the sequence $\{y_m\}$, where $y_m = e - |x_n - x_m|$, is a positive sequence in $\mathbf{R}$.

$\therefore \qquad [y_m] > 0$ in $\mathbf{R}$

or $\qquad e - [|x_n - x_m|] > 0$ in $\mathbf{R}$

Thus for $n \geq m_0$

$$\begin{aligned} |x_n - \xi| &= |x_n - [x_m]| \\ &= |[x_n] - [x_m]| \\ &= |[x_n - x_m]| = [|x_n - x_m|] < e < \varepsilon \end{aligned}$$

$\therefore \qquad \lim x_n = \xi$, and $\xi \in \mathbf{R}$.

Corollary 1. *If $\xi \in$* $\mathbf{R}$ *and $\varepsilon > 0$ in* $\mathbf{R}$ *then there is an $x \in$* $\mathbf{Q}$ *such that $|\xi - x| < \varepsilon$* in $\mathbf{R}$.

Let $\{x_n\}$ be a Cauchy sequence in $\mathbf{Q}$ such that $\{x_n\} \in \xi$.

$\therefore \qquad \lim x_n = \xi$ in $\mathbf{R}$

Hence for every $\varepsilon > 0$ in $\mathbf{R}$, $\exists$ an $m \in \mathbf{N}$ such that

$$|\xi - x_n| < \varepsilon \text{ in } \mathbf{R} \quad \forall\, n \geq m$$

In particular for $x = x_m \in \mathbf{Q}$,

$$|\xi - x| < \varepsilon$$

Corollary 2. *If $\xi < \eta$ in* $\mathbf{R}$, *then there is an $x \in$* $\mathbf{Q}$ *such that $\xi < x < \eta$.*

We know that every ordered field is dense. Therefore, there is a real number ζ such that $\xi < \zeta < \eta$. If $\varepsilon = \min(\zeta - \xi, \eta - \zeta)$, then by Cor. 1. There is a rational number x such that

$$\xi < \zeta - \varepsilon < x < \zeta + \varepsilon < \eta.$$

Corollary 3. $\mathbf{R}$ *is Archimedean*

or

For each pair of positive real numbers ξ, η, there exists a positive integer n such that $n\ \xi > \eta$.

For $0 < \xi < \eta$ in $\mathbf{R}$, let x, y be rational numbers such that $0 < x < \xi < \eta < y$ in $\mathbf{R}$. Since the field $\mathbf{Q}$ is Archimedean, therefore $\exists$ an $n \in \mathbf{N}$ such that $nx > y$.

$\therefore \qquad n\xi > nx > y > \eta$

Hence **R** is Archimedean.

Theorem 14. *Every Cauchy sequence of Cantor real numbers converges in* **R**.

Let $\{\xi_n\}$ be a Cauchy sequence in **R**. Then for $\varepsilon > 0$ in **R**, $\exists\, m_1 \in \mathbf{N}$ such that

$$|\xi_n - \xi_m| < \varepsilon/3 \qquad \forall\, m, n \geq m_1 \qquad \text{...(1)}$$

By Cor. 1, for each $n \in \mathbf{N}$, there exists a rational number x_n such that $|\xi_n - x_n| < 1/n$.

For above $\varepsilon > 0$, we can choose $m_2 \in \mathbf{N}$ such that

$$1/n < \tfrac{1}{3}\varepsilon \text{ and } |\xi_n - x_n| < 1/n < \tfrac{1}{3}\varepsilon \qquad \forall\, n \geq m_2 \qquad \text{...(2)}$$

Let $m_0 = \max(m_1, m_2)$, from (1) and (2),

$$\begin{aligned} |x_n - x_m| &= |x_n - \xi_n + \xi_n - \xi_m + \xi_m - x_m| \\ &\leq |x_n - \xi_n| + |\xi_n - \xi_m| + |\xi_m - x_m| \\ &< \tfrac{1}{3}\varepsilon + \tfrac{1}{3}\varepsilon + \tfrac{1}{3}\varepsilon = \varepsilon, \quad \forall\, m, n \geq m_0 \end{aligned}$$

$\therefore$ $\{x_n\}$ is a Cauchy sequence in **Q**.

Hence $[x_n]$ is a real number ξ, say,

$\therefore$ $\lim x_n = \xi$ in **R** [Theorem 13]

Again, since $\{x_n\}$ converges to ξ in **R**, for $\varepsilon > 0$ in **R**, $\exists$ an $m_3 \in \mathbf{N}$ such that

$$|x_n - \xi| < \frac{2}{3}\varepsilon, \quad \forall n \geq m_3 \qquad \forall\, n \geq m_3 \qquad \text{...(3)}$$

Hence for $n \geq m' = \max(m_2, m_3)$, we have

$$\begin{aligned} |\xi_n - \xi| &= |\xi_n - x_n + x_n - \xi| \\ &\leq |\xi_n - x_n| + |x_n - \xi| \\ &< \tfrac{1}{3}\varepsilon + \tfrac{2}{3}\varepsilon = \varepsilon \qquad \text{[using (2), (3)]} \end{aligned}$$

$\therefore \qquad \lim \xi_n = \xi$ in **R** .

7. COMPLETENESS IN R

Theorem 15. ***Order-completeness property.*** *Every non-empty subset of real numbers which is bounded above has the supremum in* **R**.

(*i*) Suppose A is a non-empty subset of **R**, b is an upper bound of A, and $\alpha \in A$ so that $\alpha \leq b$.

Since **R** is Archimedean, therefore, for each $n \in \mathbf{N}$, there exists an $\overline{m} \in \mathbf{N}$ such that $\alpha + (\overline{m}/n) \geq b$ in A, and therefore $\alpha + (\overline{m}/n)$ is an upper bound of A. Hence for each $n \in \mathbf{N}$, the set

$$B_n = \left\{ m \mid \alpha + \frac{m}{n} \text{ is an upper bound of } A, m \in \mathbf{N} \right\}$$

is a non-empty subset of **N**.

But, since every non-empty subset of natural numbers has a least element, let m_n be the least element of B_n. Thus for each $n \in \mathbf{N}$,

(1) $y_n = \alpha + \dfrac{m_n}{n}$ is an upper bound of A, and

(2) $x_n = y_n - \dfrac{1}{n} = \alpha + \dfrac{m_n - 1}{n}$ is not an upper bound of A.

$$\therefore \qquad x_m < y_n, \quad \forall\, m, n \in \mathbf{N}$$

Now

$$x_n - x_m < y_m - x_m = \frac{1}{m}$$

and

$$x_m - x_n < y_n - x_n = \frac{1}{n}$$

$$\begin{aligned}\therefore \qquad |x_n - x_m| &= \max\,(x_n - x_m, x_m - x_n)\\ &< \max\left(\frac{1}{m}, \frac{1}{n}\right), \quad \forall\, m, n \in \mathbf{N}\\ &< \frac{1}{n_0}, \text{ for } \; m, n \geq n_0\end{aligned}$$

Hence, $\{x_n\}$ is a Cauchy sequence in $\mathbf{R}$ which by Cauchy property of $\mathbf{R}$ (Theorem 14, § 6) converges in $\mathbf{R}$.

$$\therefore \qquad \lim x_n = \xi, \text{ where } \xi \in \mathbf{R}.$$

(*ii*) We shall now show that $\xi = \sup A$.

Let, if possible, ξ be not an upper bound of A.

Hence $\xi < x$, for some x in A.

Since $\lim x_n = \xi$ and $\lim 1/n = 0$, there is some $n \in \mathbf{N}$ such that

$$\frac{1}{n} < \frac{x - \xi}{2}$$

and

$$x_n - \xi \leq |x_n - \xi| < \frac{x - \xi}{2}$$

Then by equations (1) and (2), we have

$$y_n = x_n + \frac{1}{n} < \left(\xi + \frac{x - \xi}{2}\right) + \frac{x - \xi}{2} = x \text{ in } \mathbf{R}$$

But this is impossible by (1), since $x \in A$.

Hence ξ is an upper bound of A.

If ξ is not the least upper bound, let $\eta < \xi$ be the least upper bound of A.

Let $\xi - \eta = \delta > 0$.

Since $\lim x_n = \xi$, therefore for $\delta > 0, \exists\, n \in \mathbf{N}$ such that

$$\xi - x_n \leq |\xi - x_n| < \delta = \xi - \eta$$

$$\therefore \qquad \eta < x_n \leq x, \text{ for some } x \in A$$

which is a contradiction.

Hence $\xi \leq \eta$, so that ξ is the least upper bound of A.

Thus, the non-empty set A of real numbers has the supremum in $\mathbf{R}$.

Bibliography

Anderson, J.A., *Real Analysis,* Logo Press Ltd., London, 1969.

Apostol, T.M., *Mathematical Analysis,* Addison-Wesley, 1974.

Apostol, T.M., *Calculus,* Vol. 1, 2nd ed., Zerox Waltham, 1967.

Asplund, E., and Bungart, L., *A First Course in Integration,* Holt Rinehart and Winston, New York, 1966.

Bak, T.A. and Lichtenberg, J., *Functions of One and Several Real Variables,* Vol. 2. Mathematics for Scientists, W.A. Benjamin Inc., New York, 1966.

Barra, G.de., *Measure Theory and Integration,* Ellis Horwood Ltd., 1981.

Bartle, R.G., *The Elements of Real Analysis,* 2nd ed., John Wiley & Sons, 1976.

Bartle, R.G. and Sherbert, D.R., *Introduction to Real Analysis,* John Wiley & Sons, 1982.

Berberian, S.K., *Measure and Integration,* Chelsa Publishing Company, Bronx, New York, 1965.

Berman, G.N., *A Problem Book in Mathematical Analysis* (Translation), MIR Publishers, Moscow, 1980.

Binmore, K.G., *Mathematical Analysis: A Straightforward Approach,* Cambridge University Press, 1982.

Bromwich, T.J.I' A. and MacRobert, T.M., *An Introduction to Theory of Infinite Series,* Macmillan & Company Ltd., London and St. Martin's Press, New York, 1959.

Burkill, J.C., *The Lebesgue Integral,* Cambridge University Press, 1951.

Burkill, J.C., *A First Course in Mathematical Analysis,* Cambridge University Press, 1970.

Burkill, J.C. and Burkill, H., *A Second Course in Mathematical Analysis,* Cambridge Univ. Press, 1970.

Burril, C.W., and Knusden, J.R., *Real Variables,* Holt, Rinehart and Winston Inc., New York, 1969.

Chae, S.B., *Lebesgue Integration,* Marcel Dekker Inc., New York, 1980.

Cohen, L., and Ehrlich, G., *The Structure of the Real Number System,* Van-Nostrand, 1963.

Demidovich (Ed.), B., *Problems in Mathematical Analysis,* MIR Publishers, Moscow, Fifth printing, 1981.

DePree, J. and Swartz. C., *Introduction to Real Analysis,* John Wiley & Sons, Inc., New York, 1988.

Devinatz, A., *Advanced Calculus,* Holt, Rinehart and Winston, New York, 1968.

Ferrar, W.L., *A Text Book of Convergence,* Clarendon Press, Oxford, 1938.

Fischer, E., *Intermediate Real Analysis,* Springer-Verlag, New York, Inc., 1983.

Folland, G.B., *Real Analysis,* Wiley Interscience, 1984.

Fulks, W., *Advanced Calculus,* John Wiley & Sons, 1978.

Gaskill, H.S. and Narayananswami, P.P., *Foundations of Analysis,* Harper & Row Publishers, New York, 1989.

Gelbaum, B.R. and Olmsted, J.M.H., *Counter examples in Analysis,* Holden-Day Inc., 1964.

Gibson, G.A., *Advanced Calculus,* Macmillan & Co. Ltd., New York, 1958.

Goffman, C., *Introduction to Real Analysis*, Harper & Row International Edition, 1967.

Goldberg, R.R., *Methods of Real Analysis*, John Wiley & Sons, 1976.

Halmos, P.R., *Measure Theory,* Van Nostrand, Princeton, N.J., 1950.

Hardy, G.H., *A Course in Pure Mathematics* (10th ed.), Cambridge Univ. Press, 1952.

Hartman, S., Mikusinski, J., *The Theory of Lebesgue Measure and Integration,* Pergamon Press, Oxford, 1961.

Hawkins, T., *Lebesgue Theory of Integration—Its Origin and Development,* Univ. of Wisconsin Press, Madison, 1970.

Hayes, C.A. Jr., *Concepts of Real Analysis,* John Wiley & Sons, New York, 1964.

Hewitt, E., *Theory of Functions of a Real Variable,* Holt Rinehart and Winston. New York (Preliminary ed.), 1960.

Hilderbrand, F.B., *Advanced Calculus for Applications,* Prentice-Hall Inc., New Jersey, 1962.

Hu, T.S., *Elements of Real Analysis,* Holden Day, Inc., 1967.

Jeffery, R.L., *The Theory of Functions of a Real Variable,* 2nd edition, Univ. of Toronto Press, Toronto, 1953.

Johnsonbaugh, R. and Pfaffenberger, W.E., *Foundations of Mathematical Analysis,* Marcel Dekker Inc., 1981.

Klambauer, G., *Mathematical Analysis.* Marcel Dekker, New York, 1975.

Knopp, K., *Infinite Sequences and Series,* Dover Publications, Inc., New York, 1956.

Knopp, K., *Theory and Applications of Infinite Series,* 2nd edition, Hafner, New York, 1951.

Kolmogorov, A.N. and Fomin, S.V., *Introductory Real Analysis.* Prentice-Hall, Inc., 1970.

Kuller, R.G., *Topics in Modern Analysis,* Prentice-Hall, Inc., 1969.

Kuratowski, K., *Introduction to Calculus* (Translated from Polish), 2nd edition, Pergamon Press, 1969.

Lang, S., *Analysis I, II,* Addison-Wesley, 1968.

Lang, S., *A Second Course in Calculus,* 2nd edition, Addison Wesley, 1969.

McShane, E.J., *Integration,* Oxford Univ. Press, London (fifth printing), 1961.

Narayan, S., *A Course of Mathematical Analysis,* S. Chand & Company, Delhi, 1958.

Natanson, I.P., *Theory of Functions of Real Variables* (Translated by L.F. Boron), Vol. 1, Frederic Unger Pub. Co., New York, 1955.

Nikolsky, S.M., *A Course of Mathematical Analysis* (Translated from Russian by V.M. Volosov), Vol. I, II, MIR Publishers, Moscow, 1977.

Olmstead, J.M.H., *Advanced Calculus,* Prentice-Hall, 1961.

Parzynski, W.R. and Zipse, P.W., *Introduction to Mathematical Analysis,* McGraw-Hill International Book Co., 1982.

Piskunov, N., *Differential and Integral Calculus,* MIR Publishers, Moscow, 1969.

Phillips, E.G., *A Course of Analysis,* Cambridge Univ. Press, 1960.

Pólya, G. and Szegö, G., *Problems and Theorems in Analysis.* Vol. 1, Narosa Pub. House, New Delhi, 1979.

Protter, M.H. and Morrey, C.B., Jr., *Modern Mathematical Analysis,* Addison Wesley, Fourth printing, 1972.

Randolph, John, F., *Basic Real and Abstract Analysis,* Academic Press, 1968.

Ray, W.O., *Real Analysis,* Prentice-Hall, 1988.

Riesz, F. and Sz-Nagy, B., *Functional Analysis* (Translation), Prederic Unger Pub. Co., New York, 1955.

Ross, K.A., *Elementary Analysis: The Theory of Calculus,* Springer-Verlag, 1980.

Royden, H.L., *Real Analysis,* 3rd edition, Macmillan, New York, 1988.

Rudin, W., *Principles of Mathematical Analysis,* 3rd ed., McGraw-Hill, New York, 1976.

Rudin, W., *Real and Complex Analysis* (2nd edition), Tata McGraw-Hill, New Delhi, 1981.

Saxena, S.C. and Shah, S.M., *Introduction to Real Variable Theory,* Prentice Hall of India Ltd., 1987.

Simmons, G.F., *Introduction to Topology and Modern Analysis,* McGraw-Hill Book Company Inc. 1963.

Spiegel, M.R., '*SCHAUM'S Outline of Theory and Problems of Real Variables* (*Lebesgue Measure and Integration*),' McGraw-Hill Book Company, Inc., New York, 1969.

Spiegel, M.R., '*SCHAUM'S Outline of Theory and Problems of Advanced Calculus*,' SCHAUM Pub. Co., New York, 1963.

Stirling, David, S.G., *Mathematical Analysis,* Ellis Horwood Ltd., 1987.

Taylor, A.E., *General Theory of Functions and Integration,* Blaisdell, New York, 1965.

Titchmarsh, E.C., *The Theory of Functions* (2nd edition), English Language Book Society, Oxford Univ. Press, 1939.

Tolstov, G.P., *Fourier Series* (Translation), Dover Publications, Inc., 1962.

Whittaker, E.T. and Watson, G.N., *Course of Modern Analysis* (4th edition), Cambridge Univ. Press, 1958.

White, A.J., *Real Analysis: An Introduction,* Addison-Wesley Publishing Company, 1968.

Widder, D.V., *Advanced Calculus* (2nd edition), Prentice-Hall of India Ltd., New Delhi, 1964.

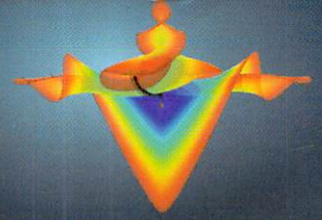
Williamson, J.H., *Lebesgue Integration,* Holt Rinehart and Winston, New York, 1962.

Wilson, E.B., *Advanced Calculus,* Dover Publications Inc., New York, 1911.

Zaanen, A.C., *Integration,* North Holland, Amsterdam, 1967.

Index

INDEX



MATHEMATICAL ANALYSIS

Contents:

The book is intended to serve as a text in Mathematical Analysis for the undergraduate and postgraduate students of various universities. Professionals will also find this book useful.

The book has theory from its very beginning. The foundations have been laid very carefully and the treatment is rigorous based on modern lines. It opens with a brief outline of the essential properties of rational numbers and using Dedekind's cut, the properties of real numbers are also established. This foundation supports the subsequent chapters: *Topological Framework Real Sequences and Series, Continuity, Differentiation, Functions of Several Variables, Elementary and Implicit Functions, Riemann and Riemann-Stieltjes Integrals, Lebesgue Integrals, Surface, Double and Triple Integrals* which are discussed in detail. Uniform Convergence, Power Series, Fourier Series, Improper Integrals have been presented in as simple and lucid manner as possible. Number of solved examples to illustrate various types have also been included.

As per need, in the present atmosphere, a chapter on Metric Spaces discussing completeness, compactness and connectedness of the spaces has been added. Finally, two appendices discussing Beta-Gamma functions, and Cantor's theory of real numbers, add glory to the contents of the book.

S C Malik was a Senior Lecturer in the Department of Mathematics, SGTB Khalsa College, University of Delhi, Delhi. He had taught graduate and undergraduate analysis courses for more than three decades.

Savita Arora was a Senior Lecturer in the Department of Mathematics, SGTB Khalsa College, University of Delhi, Delhi. She had taught graduate and undergraduate Analysis Courses for more than two decades.

Price: ₹ 499.00

ISBN: 978-93-859-2386-9

9 789385 923869

NEW AGE TECHNO PRESS

An Imprint of

NEW AGE INTERNATIONAL (P) LIMITED, PUBLISHERS

LONDON • NEW DELHI • NAIROBI

Bangalore • Chennai • Cochin • Guwahati • Hyderabad • Kolkata • Lucknow • Mumbai

visit us at **www.newagepublishers.com**

This edition is available for sale in India, Bangladesh, Bhutan, Maldives, Nepal, Pakistan and Sri Lanka only.